

More crisis for UK research

Now it seems that Britain will stay in high-energy physics for the time being, high-energy physics has an obligation to mitigate the squeeze on general research.

THE Kendrew committee, which has been looking into British membership of CERN, the European nuclear physics laboratory at Geneva, seems about to create a dilemma for everybody in sight — the British Science and Engineering Research Council, which pays the British subscription, the Advisory Board for the Research Councils, which has just presented the British government with a demand for extra cash, and even the management of CERN, the essence of the committee's conclusion seems to be that Britain should not immediately pull out of the European high-energy physics club, but instead demand that the cost of membership should be reduced from the end of this decade, when CERN will be equipping with superconducting magnets the electron collider now being built. Precisely what the committee thinks should happen if these negotiations come to nothing should be clear when its report is published. It will be particularly valuable to know why the committee has decided against withdrawal. Is high-energy physics after all worthwhile, or is it merely that the British research community will need the best part of five years to win what it can from past investment in CERN?

That the outcome is a compromise will be an embarrassment. Although high-energy physics is a steadily shrinking part of the budget of the Science and Engineering Research Council, it is still uncomfortably large. Last year, the advisory board was saying that the research council might have to withdraw from some whole field of science if the squeeze persisted. The Kendrew committee sprang from the sense that high-energy physics, the most conspicuous target for economizers, is too important to be abandoned just like that. The chief consequence of the Kendrew compromise is that there is no immediate prospect of relief. The reappraisal of all activities that has occupied the research council for the past six months will have to continue. Nor will there be much comfort in the recent guileful letter to Kendrew from Dr Jeremy Bray, the Labour Party spokesman on science and technology, pointing out that a future government might have different views about science budgets.

Speculations prompted by that kind of prospect are neither here nor there. Any British government would find itself in the same plight on public spending as does the present. Extra money for research, such as the advisory board has now asked for (£10 million extra next year and in each of the two succeeding years) would have to come from the British government's present spending programme. It would have been more interesting, but dangerous, if the Kendrew committee had resolved flatly to say that the research community cannot afford high-energy physics out of the present budget, and had thrown the onus of grappling

with the consequences onto the government.

The advisory board, whose role is to divide the science budget each year, will now be particularly under stress, if only because some of its members are at best agnostic on high-energy physics. But continued membership of CERN squares badly with economies being made elsewhere. Out of a total budget of £631 million, the research councils will have to spend £9 million this year and £11 million next year on persuading research people to retire early. That is simple waste. The advisory board is also acutely aware, to judge from this year's submission to the government, that its own economies will have to take effect before there can be much change in the pattern of spending within the universities on the support of research, which is estimated (a little optimistically) to be worth £550 million a year. Only in the academic year beginning in 1986 will there be the first signs of how the budgets of individual universities will be changed to take account of the importance attached to the research they do. So the muttering has begun that part of the university budget should be handed over to the research councils, which is a recipe for setting academics at each others' throats.

This is where the physics community, and the management of CERN, can help. Both have in the past been dealt with generously, and may by now have earned an obligation to do something for fields more recently made vigorous. The truth about CERN is that it is every bit as splendid a laboratory as its founders intended. It is a model of international collaboration, and it is successful in the work it does. For the past decade, conscious of its cost, the laboratory has also managed to keep ahead technically by finding the money with which to build a new accelerator by closing down an older machine, or by using it as a component for a larger one. But the laboratory has some defects. It has a larger permanent staff than might otherwise be necessary, perhaps because it is international. Much worse, it happens to be in Switzerland, one of the most prosperous countries in Western Europe and that which determines what people working there are paid. Nobody suggests that the new electron accelerator should be moved somewhere else (which would be impossible) or that people working at CERN should be underpaid, but this privileged laboratory should now help solve the problems that have arisen in Britain, and even elsewhere.

What can CERN do? Economizing is an obvious must, but cheeseparing would not save anything like as much as Kendrew appears to be looking for. Broadening the base of collaboration on the second phase of the electron collider is a much better course to follow. Neither the United States nor the Soviet Union is at present a paying member of the club, even though people from each country work on CERN collaborations by a kind of gentlemen's agreement, on the ground that reciprocity eventually accrues. Why not now formalize these arrangements? If CERN has shown that European collaboration in high-energy physics works, why should it not use its present position on the crest of a wave of success to show that international collaboration can also succeed? With the US Congress in the thick of trying to reduce the US budget deficit, this might be just the time for such a move. One of the curious ironies of this field, over several generations of particle accelerators, is that the practitioners are forever saying that machines will eventually have to be built and operated internationally — but not just yet.

Biological manuscripts

From the beginning of next week (20 May), Miranda Robertson, Biological Sciences Editor of *Nature*, will be based in the Washington office of *Nature*. Manuscripts offered for publication may be sent, as at present, to either the London or the Washington office; with the benefit of good communications, the speed with which manuscripts are dealt will be independent of the place at which they are received. But authors are asked please in future to send **four copies of manuscripts** intended for publication (one for each office and two for referees). □



US universities

Task force quests for financial wizardry

Washington

THE Congress's Science Policy Task Force was told last week that between \$5,000 million and \$20,000 million might be needed to bring US university research facilities up to acceptable modern standards, and that the existing partnership between universities, industry and government "may not be adequate" for the task. While the general sentiment has been heard before on Capitol Hill, this particular warning was surprising because it was uttered not by some university president making a cap-in-hand appeal for a new laboratory, but an assistant director of the White House's Office of Science and Technology Policy, Dr Bernadine Healy.

In two days of hearings on "government and the research infrastructure", the task force heard a variety of prescriptions for improving universities' access to state-of-the-art equipment. In Healy's view, both government and universities must take their share of the blame for the accumulated deficit of research hardware. Universities, she said, have often behaved like dependants of the government, "abdicating their responsibility for infrastructure and biding their time until federal facilities and programmes were resumed". The government, for its part, has "attempted not to invest in the research enterprise, but to procure packets of research results at the lowest possible prices".

The Science Policy Task Force, a bipartisan group established within the House of Representatives' Committee on Science and Technology, has been hearing from witnesses in government and academic life since mid-April. The hope is that a preliminary report will be completed by May next year, with a final version available the following October. Witnesses so far have fallen into two clearly defined camps: those who believe that the present system for provision of research facilities is working well, and those who feel that catastrophe is imminent.

The latter group argues that universities grew accustomed to 15 per cent annual increases in their operating research budgets during the heady 1970s, and even during the early years of the first Reagan presidency. As long as the hefty increases continued, universities were cushioned from the effects of the lack of investment in buildings and large items of equipment: most of the federal programmes specifically geared to improving research infrastructure had dried up in the early 1970s. Now that the days of 15 per cent annual increases are seemingly over, the price of years of underinvestment in the infrastructure will have to be paid.

Healy, speaking with the advantage of inside knowledge of the White House's soon-to-be-completed study of the universities chaired by David Packard, said that a multi-billion dollar programme might improve conditions, but that changes in attitudes would also be needed to put things to rights. Specific proposals include the following:

- the proportion of the \$20,000 million civilian research and development budget spent in universities should be increased above its present level of 30 per cent;
- unrestricted donations of equipment and contributions to renovations should be encouraged (presumably by favourable tax exemptions);
- amortization periods for both equipment and buildings assumed in federal grants should be reduced from their present unrealistically high levels (50 years and 15 years respectively) to something closer to reality (say, 20 years and 6-8 years);
- and (as a strong hint) the government should stop trying to "micro-manage" equipment purchases and reduce the burden of excessive documentation.

The need for innovative financial approaches was echoed by Dale Corson, chairman of the government/university/industry research round table sponsored by the National Academies of Science and Engineering. Corson asserted that obsolescence of research instruments is limiting productivity, and that in engineering, in particular, this state of affairs is increasingly driving away potential recruits into the field.

Corson argued that the problem has arisen largely because of the increased cost of research instruments; he proposed that more use should be made of shared facilities and that more facilities should be financed jointly with state governments. Besides exploring new ways of providing equity, the traditional form of financing for research facilities, universities should also look at new ways of debt financing. Corson asked for the removal of legal obstacles that, in essence, prevent the promise of federal grants from being used as collateral against loans.

Summing up in a virtuoso exhibition of Washington bureaucratese, Corson asked that these initiatives be brought together in a national programme that will "regularize the facilities appropriation process" and will "leverage federal funds to the maximum degree possible". Corson's proposal will doubtless be aired in July at a conference on academic research facilities sponsored by the government/university/industry research round table and federal agencies.

Tim Beardsley

Growth hormone

FDA ban on pituitary product

Washington

FEARS that growth hormone extracted from human pituitary glands may spread the rare neurological condition known as Creutzfeldt-Jakob disease (CJD) have prompted the US Food and Drug Administration (FDA) to stop the use of all human pituitary products. Those most immediately affected are the 2,500 hypopituitary dwarfs in the United States, who depend on the hormone to maintain normal rates of growth; clinical research on other pituitary products has also been suspended. FDA is now under strong pressure to approve the general use of a recombinant human growth hormone product manufactured by Genentech Inc. that has been under FDA review for the past 18 months.

The two commercial companies that have now withdrawn their human-derived growth hormone products, KabiVitrum and Serano Laboratories, point out that the evidence against them is at best indirect. There have been three recent deaths attributed to CJD; among young men who received crude pituitary extracts as children; suspicions were aroused because the disease is normally found only in those over 40. Only one diagnosis has been confirmed, however, and no autopsy was carried out on one of the suspected cases. In the other two cases, there are other complicating factors that could explain infection by a slow-acting virus such as that believed to cause CJD; one was a diabetic who received frequent insulin injections, for example.

The deaths occurred among patients who had been treated under the National Hormone and Pituitary Program of the National Institutes of Health (NIH). Since the programme started, the purification methods used have improved markedly, and all growth hormone manufactured since 1977 has been through a final stage of column chromatography using Sephadex. The commercial manufacturers, who also use column chromatography, point to a study by Professor A.G. Dickinson of the University of Edinburgh which indicated that virus deliberately introduced into pituitary tissue was not detectable after the column chromatography stage of purification.

One argument against a link between CJD and growth hormone is that the hormone is produced in batches that are used to treat several hundred children. Even if only one batch had been contaminated, therefore, many more than three cases would be expected. There are not known to be any living growth hormone recipients with CJD symptoms, and a recent survey of 300 recipients in Switzerland found no cases of CJD.

Erol Caglarcan of Serano accuses FDA of using "unnecessary pressure" and says the Serano product, which has the trade name Asellacrin, will still be used outside the United States. Serano hopes in time to market a genetically engineered growth hormone being developed by Celltech in England, but that product is not yet at the stage of clinical trials. KabiVitrum's recombinant product, developed under licence from Genentech, is already in clinical trials in Europe: the company is correspondingly less concerned about FDA's action in banning human growth hormone.

The pressure on FDA to approve Genentech's recombinant product as quickly as is seemly arises because testing the pituitary samples for presence of the CJD infective agent has to be done by bioassay, using squirrel monkeys or chimpanzees. Besides being expensive, results of tests now being mounted by NIH will not be available for two years. In any case, negative results from such tests would not provide much reassurance, given that so little is known about the disease. If Genentech's product can be shown to be safe, FDA might be tempted to abandon pituitary-derived growth hormone permanently.

A decision to abandon pituitary products would, however, affect other substances besides growth hormone. Dr A.F. Parlow of the University of California at Los Angeles, who has supplied pituitary products to NIH since 1977, is concerned that research into other pituitary hormones could be set back; clinical trials of prolactin and thyroid-stimulating hormone, for example, have had to be put on ice. Parlow believes that FDA will have to embark on a convincing demonstration that viruses can be effectively excluded by modern purification methods in order to prevent a complete halt in pituitary research.

Genentech first produced a recombinant growth hormone product several years ago, but — perhaps because of the ready availability of the human-derived product — progress towards FDA approval has been slow. The Genentech hormone has an extra methionine residue tagged on that is not found in the natural substance, and in early clinical trials many patients developed antibodies to the product.

Last year, FDA asked Genentech to provide further safety data on a trial population to be monitored for a full year; that study is now under way but the data requested by FDA will not be complete until the start of 1986 at the earliest. Recently, however, FDA officials have been saying publicly that the Genentech product, which has now been improved, is likely to be approved within the next two or three months.

Genentech, for its part, is playing it cool, saying the decision is entirely up to FDA; a recombinant growth hormone without the extra methionine is also believed to be in the pipeline.

Tim Beardsley

French science budget

Fabius goes for growth

M. HUBERT Curien, French minister for research and technology, at last has his figures. For the past few weeks he has been waiting to hear from the Prime Minister, M. Laurent Fabius, how much money there is in the kitty for French research for the next three years, to put some substance in his "three-year plan" for research — an instrument which parliament must vote on and which should give a new impetus to the political emphasis on science in France. The Prime Minister's answer is four per cent real growth per year in civil research and development spending to 1988, excluding defence research spending.

And that is just the government side. There will also be increased incentives to industry to carry out research (a doubling of tax allowances on research and development budgets) which will give industry another FF 600-700 million (£60-70 million) a year to spend, so that overall it is estimated that French national research and development spending will rise over the next three years from 2.25 per cent of gross national product now to 2.6 per cent in 1988. At its nadir in 1979-80, this fraction had touched 1.8 per cent.

Thus the growth in French research spending should continue, provided there are no global budget "corrections" in future, corrections which rather reduced the true impact of the ambitious "law for science" introduced in 1982 by M. Curien's predecessor, M. Jean-Pierre Chevènement. But Chevènement also increased the number of jobs for scientists, and this trend is to continue. While British research councils have lost around a fifth of their posts since 1981, their French counterparts have gained something around 3 per cent a year. Curien's plan offers 1,400 new jobs for scientists over the next three years, less annually than last year's 1,000, but still a substantial figure.

So where are the French after all this attention to science? Doing fairly well, but could do better, particularly in industry, is the informed verdict in Paris.

Briefly, the tally is thought to be this. First, a reversal of the decline of support for French science in the 1970s. The decline was brought about initially by President Pompidou's massive indifference, and continued by successive administrations until Pierre Aigrain, science minister under President Giscard d'Estaing, reversed the trend. Second, M. Chevènement, under President Mitterrand, made science a political touchstone, and changed the whole French political mentality towards science, and of scientists towards industry. Third, the research councils — such as the Centre National de la Recherche Scientifique (CNRS) with its 10,000 scientists — were accorded looser legal ties, which allowed them to make profitable associations with industry, set up affiliates and so

on. And fourth, science and technology were given weight, and budgets, in the regional administrations of France, which gave an opportunity for a more flexible response to local industrial and social needs. A fifth bonus, the new impetus given to universities to compete with one another for students and research cash, has come too late to yet have any real impact.

Failures, however, include a poor response from industry, too much centralization of decision-making and too little realistic assessment of the successes and failures of the policy.

There is hope, however, that Curien will be able to correct many of these failings. The new three-year plan, agreed at ministerial level but not yet voted upon, is one of his instruments. A practical man, Curien seems to be taking many of the right steps to put the rebirth of French science back on course.

Robert Walgate

Australia's NDP in disarray

Canberra

AUSTRALIA'S single-issue Nuclear Disarmament Party (NDP), whose 8,000 members include many former Labor Party supporters disaffected by the Hawke Labor government's decision in June last year to mine and export uranium, has lost its only parliamentary representative-elect, Ms Jo Vallentine, who resigned last week, announcing that she would set up a new anti-nuclear political party to be known as Peace and Nuclear Disarmament Action (PANDA).

This move follows NDP's first national conference in Melbourne last month in which Senator-elect Vallentine, former Labor Senator Ms Jean Melzer and narrowly-defeated NDP Senate candidate and rock singer, Mr Peter Garrett, staged a walkout, claiming that NDP was in danger of falling under the domination of entryist members of the Socialist Workers Party, a Marxist grouping of Trotskyite tendency. The conference was to have given some direction to NDP policy, because the party had only six months to organize before last December's half-Senate elections.

One of the chief reasons given by the breakaway group for leaving NDP was the party's "infiltration" by people who were already members of other political parties: their proscription is likely to be a feature of membership of PANDA. The members of Ms Vallentine's West Australian branch of NDP were canvassed by a postal ballot that ended in a vote to cut all ties with the national body and form a separate party. Breakaway groups in most other states are holding similar ballots whose results will be known next month.

Jeffrey Sellar

US investment

Universities and South Africa

Washington

THIS spring has brought to many US college campuses demonstrations, reminiscent of the 1960s, against apartheid in South Africa. With the demonstrations has come a reconsideration of an old issue: should universities invest in companies that do business in South Africa?

According to several recent surveys, fewer than a dozen universities have completely divested themselves of holdings in such companies. And even as pressure mounts to divest, a new study by the Investor Responsibility Research Center, an organization founded 13 years ago to provide institutional investors with information on the ethical behaviour of corporations, may stiffen the resolve of those university boards that have resisted so far: according to the report, "divestment will have a detrimental effect over the long term on portfolio performance".

Most of the hundred or so universities that have taken some action on investments in South Africa, mostly in response to protests in the early 1970s, have adopted policies of partial, selective divestment, usually selling off stock in companies that fail to adhere to the so-called Sullivan principles. These rules, developed by Reverend Leon Sullivan of Philadelphia, a member of the General Motors board of directors, prescribe basic fair employment practices for companies operating in South Africa. The universities that have opted for partial divestment argue that a selective policy is the best method of influencing corporate behaviour; with across-the-board divestment, influence is lost.

Those calling for total divestment now say that this selective policy simply has not produced results. According to a recent survey by the American Council on Education (ACE), only about 100 of the 350 US companies operating in South Africa adhere to the Sullivan principles; and in any case there is no evidence that other foreign companies or South African companies have been influenced by the example of these 100. Bishop Desmond Tutu, who won the Nobel peace prize for his stand against apartheid, has called for a more stringent set of principles: companies in South Africa should recognize black labour unions, invest in black education and training, and should freely hire blacks without regard for the country's restrictive laws governing the movement and residence of the 24 million South African blacks.

The most militant protests this spring have taken place at Columbia University and the University of California at Berkeley, both of which already have policies of partial divestment. Students have occupied buildings and there have been a number of arrests.

And while both of those universities have agreed to reconsider their policies, others

are being more outspoken in their belief that the chief responsibility of university financial administrators is to make money for the university. Derek Bok, president of Harvard, which has maintained a selective divestment policy since 1981, spoke for many when he said recently: "Despite our revulsion toward apartheid, the fact remains that Harvard's resources were entrusted to us for academic purposes and not as a means for demonstrating our opposition to apartheid or to other manifest injustices and evils around the world." The

ACE survey noted that even among those universities choosing total divestment, the justification was that colleges "should not benefit from corporate practices that cause social injury"; they rejected the notion that they were taking a political stand.

US investment in South Africa is estimated at \$14,000 million. About one-third of the companies on Standard & Poor's 500-company stock index do business in South Africa.

Besides the efforts on campus, a number of states and municipalities have ordered divestment. A bill now before the US Congress would ban bank loans and computer sales to the South African government.

Stephen Budiansky

Major universities that have divested stock in companies operating in South Africa

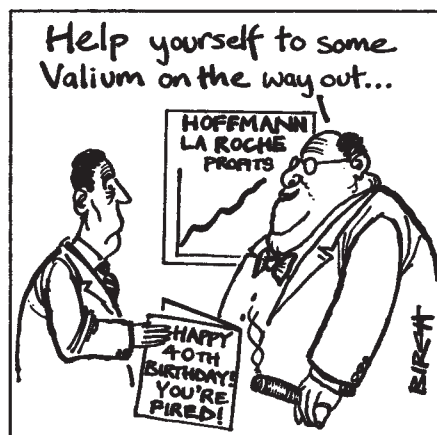
	Policy	Value of stock divested	Year action taken
Boston University	Partial	\$6.6 million	1979
Brown University	Partial	4.6	1984
City College of New York	Total	7.1	1984
Columbia University	Partial	2.7	1979
Dartmouth College	Partial	2.0	1985
Harvard University	Partial	50.9	1981
		1.0	1985
Michigan State University	Total	7.2	1979
Ohio State University	Partial	0.3	1978
Rutgers	Partial	7.0	1985
Swarthmore College	Partial	2.2	1981
Tufts University	Partial	0.1	1979
University of California, Berkeley	Partial	4.0	1979
University of Michigan	Partial	41.0	1983
University of Pennsylvania	Partial	0.8	1983
University of Wisconsin	Total	11.0	1978
Yale University	Partial	1.6	1979
		4.1	1984

Universities that have adopted a policy of total divestment have sold all holdings in companies that do business in South Africa. Partial divestment refers to a policy of selling stock in companies that have failed to adhere to the Sullivan principles or other standards of conduct or that refuse to provide information. Source: American Committee on Africa.

Age discrimination alleged

Washington

HOFFMANN-La Roche employees laid off earlier this year in what the company described as an austerity measure have



sued, alleging age discrimination. The suit, so far joined by 250 laid-off workers all over age 40, accuses Roche of pursuing a deliberate policy of singling out older employees for dismissal. More than 1,000

employees were dismissed in a move to save \$50 million a year for the next five years. Lawyers for the plaintiffs say that as many as 600 of these employees could be eligible to join the suit, which they hope to have certified as a class action.

The suit also alleges that the company's policies governing promotions, transfers and assignments favour younger workers at the expense of those over 40.

Federal law prohibits age discrimination in employment. Companies that violate the act are liable for back pay, compensation for loss of future earning power, attorneys' fees, and, if "willful" violation of the act can be shown, double damages.

Leonard Flamm, an attorney for the plaintiff, says that most of the layoffs occurred at Nutley and Belvidere, New Jersey.

Hoffmann-La Roche says that it intends to defend all such lawsuits brought against it and maintains that all decisions to terminate employment were "nondiscriminatory". No date has been set for preliminary hearings.

Stephen Budiansky

Japanese psychiatry

Abuse for visiting scientists

Tokyo

WESTERN research workers studying such topics as neurotransmitter receptors and the mode of action of antidepressants are not used to being thought of as brutal oppressors of society, nor do they conduct research in fear of assault. But, as two visiting lecturers found recently, for workers in these fields there are real dangers in Japan.

The two lecturers, Tim Crow of the Clinical Research Centre in the United Kingdom, and D. van Kammen of the University of Pittsburgh, were due to give talks on "Genes and viruses in schizophrenia" and "Episode markers in schizophrenia" at the seventh congress of the Japanese Society on Biological Psychiatry at Gifu, in central Japan. Soon after their arrival in mid-April, however, the speakers were told that the meeting had been cancelled because of threats made by "anti-psychiatrists". Instead, a smaller private meeting was arranged at a hotel in nearby Nagoya. But, twenty minutes after Dr Crow's talk began, news came that a group of 70 demonstrators were on the way to the new location; the talk was abandoned and the meeting hall evacuated.

Behind these strange events lies a 16-year feud within Tokyo University's hospital and accusations of serious abuses within Japan's psychiatric hospitals. These accusations resulted in an inspection visit earlier this month from Disabled People's International (DPI), based in Stockholm and, two weeks ago, from the International Commission of Jurists (ICJ) and the International Commission of Health Professionals (IChP), both based in Geneva.

The Gifu meeting had itself attracted attention from anti-psychiatrists because lurid publicity had just been given to the ethics of research performed at Gifu University. In February 1984, a schizophrenic woman patient admitted to a local hospital had received an abortion; the fetus's brain had then been examined in the university to look at the effects of drugs the patient had been taking. Accusations followed this year, however, that the patient had wanted to have a child and that she had been forced into the abortion in order to provide research material for a "human experiment". Although the matter was taken up by the Diet, investigations have not so far revealed malpractice.

The "anti-psychiatrists", however, see this experiment as just one more example of the "research supremism" they claim permeates the psychiatric community and which regards psychiatrically-ill patients as being without rights, mere fodder to further scientific careers.

This view has been propagated by a radical group of doctors who, in a truly bizarre episode, established themselves within Tokyo University's hospital in the

wake of the student revolt in the 1960s. At the end of the troubles, in 1969, some radical doctors, opposed to the current form of psychiatric education and treatment, seized the in-patient psychiatric ward and used violence to prevent those who did not support them from entering the ward. Since then most of the university's neuro-psychiatrists have not been able to enter the building or see the remaining patients.

An attempt was made to solve the problem in 1972. Relatives of confined patients were asked to consider moving them elsewhere; some, but not all, did so. Eventually, as doctors and nurses left, and radical doctors from elsewhere began to work in the ward, the point was reached when there were no officially-authorized staff left. The university, however, continued to supply drugs, food and equipment and to avoid confrontation on the grounds that that would not be in the interest of the remaining patients.

This strange situation persisted until 1978, when the Diet appointed an investigatory committee. The committee put pressure on the university to solve the problem without suggesting how it might be done.

The university eventually chose to avoid a fight and to legitimize the occupiers' position. University posts were given to the radical doctors, the professor was in turn allowed to make a once-a-week token visit to the ward, and a single patient was sent there from the out-patient department.

Relations between the "orthodox doctors" in the out-patient department and the "radical" doctors in the in-patient ward have not however, improved. Out-patient doctors complain of frequent abuse and physical assaults from radicals who enter the out-patient ward, often even in front of patients undergoing treatment. No evidence of unethical conduct by members of the out-patient group has been presented to justify these assaults; instead it seems as though they are being made to answer for abuses that have occurred elsewhere — the very abuses that have led to appeals to the United Nations and visits from human rights groups.

The most notorious incident occurred last year at the Utsunomiya hospital where it was alleged that patients were regularly beaten, two of them to death, one for trying to escape and one for complaining about the food, and that unqualified staff and some patients were forced to act as nurses. The director was eventually arrested, although the case has never been completely cleared up.

The official government view, echoed by many psychiatrists, is that the Utsunomiya case is a rare exception. One person who believes otherwise is lawyer Etsuro Totsuka, who for the past few years has been working almost without pay investigating

psychiatric abuses. He stresses that there is an abnormally large number of psychiatric patients in Japan (more than 300,000), many of them hospitalized for exceptionally long periods and most of them confined against their will (more than 80 per cent, against only 5 per cent in the United Kingdom). Hospitals, the great majority of them private and generating large profits, are usually built in remote areas where land is cheap but rehabilitation of patients through integration back into the community almost impossible.

Two main factors seem to lie behind these trends. First, Japan's Mental Health Law Article 33 allows a hospital administrator to commit people without their consent so long as parents or guardians agree; expert opinion from doctors is not required. Thus, as an editorial in the *Asahi* newspaper put it, there need be no surprise at the abnormally large number of people in mental hospitals because "decisions on whether or not to put the mentally ill in hospital are made, in many cases, by the hospital managers, who stand to profit from hospitalization".

The second factor is the considerable stigma still attached to the mentally ill. All too often, families with mentally-ill kin seem to want to dump them where they cannot be seen. Complaints of ill-treatment are thus likely to be few as people prefer not to be identified as having a mental patient in the family.

Detailed reports are not yet available from the three groups that made visits to Japan following lawyer Totsuka's appeal to the United Nations through the Fund for Mental Health and Human Rights that he helped set up. But James Donald of DPI has made an interim statement that "there appears to be no requirement that the decision to commit be made by a qualified expert... there is no requirement for independent review of the decision... no effective independent counsel is provided to the patient; and there appears to be no access to information on care and treatment of patients once they are put inside and consequently no protection from abuse". Lack of appropriate laws and regulations, and the absence of government supervision, lack of financial incentives to medical professionals for treatment in the community, and lack of general recognition of these rights by the larger community are all blamed for the current state of affairs.

In the eyes of foreign experts, legal reform is thus clearly indicated. Within Japan, however, the issue has so far excited little interest, except among those whose views have already become extreme. What is particularly lacking is a broad-based reform group, like MIND in the United Kingdom, even though there is a movement towards community care among psychiatrists. Even if the true scope of psychiatric abuse can be assessed and, if abuse exists, legal reforms made to remedy it, it is hard to see psychiatrists and "anti-psychiatrists" being reconciled.

Alun Anderson

Poland

University advances checked

THE Polish government's plans to reform the 1982 Higher Education Act will go ahead, in spite of strong opposition from the academic community. Last week, Warsaw University held a massive protest meeting and launched a last-minute appeal to the Sejm (Parliament) against the changes. Nevertheless, it now seems obvious that the government and party authorities were determined from the beginning to press ahead with the amendments, and the public discussion, launched at the end of December, was merely a placebo to the university community.

The amendments, which would annul the autonomy won during the Solidarity era and would give the Ministry of Science and Higher Education virtually unlimited disciplinary powers over the universities, have been strongly opposed not only by the academic community, but also by the government-sponsored "Patriotic Movement for National Rebirth" (PRON), launched during the martial law period in order (it was hoped) to attract the support of the masses to the ruling junta (WRON — the similarity of the acronyms was hardly coincidental). PRON in fact has not picked up the expected mass support, although it has consistently tried to woo the academic community, backing, for example, a proposal to convert the Białystok "filial" of Warsaw University into an independent university by 1991.

Originally, it appears, the state and party authorities had hoped to amend the law without public outcry. Unfortunately, in October 1984, a confidential draft of the proposed changes was leaked to the underground press. Although Dr Benon Miskiewicz, the Minister of Science and Higher Education, denounced the alleged leak as a "provocation", he shortly afterwards laid his own proposals for change before the Main Council for Higher Education (a kind of parliament of the universities, with one elected delegate from every higher educational institution in Poland, except the Catholic University of Lublin which elected to remain aloof when the Main Council was constituted in 1981). Although 60 per cent of the delegates to the council are party members, they strongly opposed the changes, on the grounds that to revise the law so soon after enactment would tend to "destabilize" the universities. Largely as a result of the main council's opposition, a "public debate" on the amendments was launched at the end of the year by the Committee for Social and Political Affairs, which reports directly to the cabinet.

Articles appeared in the press from leading academics condemning the amendments, and virtually the only organization that favoured them was the party-approved Association of Polish Students (ZSP), which under the proposed changes would

become the only representative organization of Polish students. (At present, ZSP has very little active support). Referenda organized among university staff and students produced almost embarrassingly high votes against the changes. Nevertheless, the usefulness of such referenda and articles was questioned by a document signed by the underground Solidarity leadership of Warsaw's eight universities and higher colleges, which, by the end of February, could be found lying about in classrooms and libraries. This claimed that the alleged public discussion was merely a means of allaying public concern, while the authorities pressed ahead with the changes.

At about the same time, a ministry spokesman admitted that the issue was simply whether or not the universities "belonged to the state", and that if they did, the changes must, willy-nilly, go through. (My tape-recording of this interview was subsequently confiscated by the

Polish customs (see *Nature* 14 March, 1985, p.120).) After the referenda ended in mid-March there were several weeks of seeming impasse. Then, on 13 May, during a special plenary meeting of the Central Committee of the Polish United Workers' Party on the role of the intelligentsia, Politburo member Jozef Czyrek stated explicitly that the universities were indeed "state owned and socialist" and that therefore the changes must go through, to ensure that "anti-socialist groups" did not exploit university autonomy for their own ends.

A week later, on 21 May, the government press spokesman, Mr Jerzy Urban, announced that there would be no further consultations with the public and that although he could give no date, the amendments would soon be laid before the Sejm. Warsaw University's 3,000-strong protest meeting the next day, which included an address by the former leader of the university's Solidarity chapter, Maciej Geller, and from which all outsiders were excluded by student "guards", made it clear that the university community will fight to preserve its hard-won autonomy. **Vera Rich**

Synchrotron source

Italy drives a hard bargain

ITALIAN science minister Luigi Granelli is not at all happy with the Franco-German *fait accompli* that led, effectively, to the choice of Grenoble in south-east France as the site for the planned European Synchrotron Radiation Facility (ESRF). Mr Granelli may have to accept the Grenoble site — Italy had proposed Trieste — but France and Germany may see little sign of his lira unless, in the accepted diplomatic tradition, Italy receives something in exchange. It now seems that he is not to be bought off with a promise that a European tritium-handling facility — needed for the time when the Joint European Torus, JET, or its successor, the Next European Torus, NET, need to use tritium to reach nuclear ignition — will be built at Ispra, north of Milan. It had been thought that this was the *quid pro quo* that would yield Italian acquiescence to, and money for the Grenoble site, but this is "not quite right", Mr Granelli said in London earlier this month.

At present there is deadlock in the European council of ministers, which meets next on 4 June, over ESRF, Granelli intimated, as "everyone is using their veto" to prevent France and Germany moving forward. Italy has proposed a site in Trieste, a region which the government is keen to develop in order to maintain its historical foothold at the top of the Adriatic, and offered more financial support than any other single country (including France or Germany). And still Italy must have something in Trieste to balance its support for Grenoble, Granelli insisted.

"There could be the large laboratory in Grenoble, and something different in Trieste", he said. "We should look at the

wider context and consider making two or three laboratories." He was prepared to be flexible about what those laboratories would be. Denmark for example, which had offered a site at Risø for ESRF, would be happy with a laboratory for oceanography.

"The council should always consider a number of facilities at the same time, then the opposition doesn't become so stiff", said Granelli. The tritium handling facility could not go into the package with ESRF, however, because "from all points of view it is not an Italian request. It's a common interest for the European Community."

Granelli would like to see big machines for Europe discussed in a more flexible way in future "as long as the Commission of the European Communities is involved". There could be arrangements where two or three countries might agree on a facility but the Commission should also be made a partner, to allow other European states to participate in a smaller way through the Commission.

Professor Paolo Fasella, director-general for research at the Commission in Brussels, agrees that this would be a useful formula. "In a slow and practical way we're already investigating it", he said. The Commission had made initial approaches to national laboratories such as the Rutherford Appleton Laboratory and the Daresbury Laboratory in the United Kingdom and Frascati in Italy, for part-participation on behalf of the Commission's "stimulation programme". "We want to develop collaboration in a flexible way", said Fasella.

Robert Walgate

Ganges pollution

Cleansing river due for cleaning

New Delhi

INDIA'S holy river Ganges, which has been washing away the sins of devout Hindus for centuries, is itself to have a clean-up. The Indian government has set a five-year deadline for the first phase of cleaning the 2,500-km river, at a cost of £180 million. The river carries 468,000 million cubic metres of water a year, a quarter of the country's surface water.

The task has been assigned to the newly formed Ganga Authority, headed by Prime Minister Rajiv Gandhi himself. The eight other members include planning and environment ministers of the central government and chief ministers of three of the eight states through which the Ganges flows. There is also a 17-member steering committee under the chairmanship of environment secretary Dr T.N. Koshoo.

Pollution of the Ganges starts at Hardwar, where it enters the plains of Uttar Pradesh, 350 km from its source in Gangotri in the Himalayas. From Hardwar to Calcutta, the whole of the river has been declared by the Central Pollution Control Board (CPCB) as unfit even for bathing, and unsuitable for drinking without massive treatment. The board says the water has become so saline and so contaminated with toxic heavy metals that some parts of the Ganges basin are now infertile and the groundwater is polluted.

The Ganga Authority aims to restore the water quality under a plan drawn up by CPCB. Dr Nilay Choudhuri, CPCB's chairman, says that 80 per cent of the pollution is caused by raw sewage discharged directly into the river from 100 cities on its banks. The rest is caused by untreated effluent, wallowing and bathing of cattle, pesticide and fertilizer run-off and incompletely cremated dead bodies.

Apart from Hardwar, the most polluted places are Benaras, Kannauj, Allahabad,

Patna, Calcutta and Kanpur, where the Ganges water is thick and turbid for 10 km because of effluent from tanneries, textile mills and chemical plants. Monkeys in Benaras have become addicted to opium in wastes discharged from a nearby factory making morphine from opium. On two occasions, the Ganges has caught fire because of oily effluent from a refinery in Barauni and discharges from a foundry forge plant in Hardwar.

Stopping the discharge of sewage into the river will reduce pollution of the Ganges by 75 per cent according to Dr Koshoo. Twenty-seven cities with populations of more than 100,000 dump 900 million litres of municipal waste each day. The plan is to renovate existing sewerage systems and pumping stations and to build new drains to divert the wastes to what are called Resource Recycling Units (RRU), where sewage will be digested to produce methane gas that will generate electricity. The residue will be sold as manure. Liquid remnants will be cleansed by biological aerators and the water produced will be used for irrigation and for growing algae and fish for sale.

According to Choudhuri, each RRU serving a population of 100,000 is expected to make an annual profit of £850,000, a sufficient incentive for private companies to take over the operation of the units after they have been built with government funds. In the second phase, the plan will be extended to the remaining 63 cities, and community cattle-sheds will be set up to collect animal wastes for processing at RRUs.

Unlike the Thames or the Danube, the cleaning up of the river Ganges is complicated by the fact that it is a sacred river to the Hindus. Millions of pilgrims immerse themselves in the river during festivals such as the Kumbha Mela and return home with a vessel filled with Ganges water.

Choudhuri says that a mobile laboratory found 10,500 coliform counts in 100-ml samples taking during a Kumba Mela in Allahabad. In Hardwar, 17.5 million litres of human and industrial waste pour daily from five drains, one of which is just upstream from where 10,000 people bathe every day and a million bathe during the Kumba Mela. At Benaras, 70,000 bathe daily in water that has been turned into an open sewer with the outfall of 120 million litres a day of sewage from 66 drains. Cremation in Benaras is the wish of every Hindu, so 400 half-burnt bodies float in the river on any day, and 9,000 cattle are thrown in each year. It is intended that electric crematoria will be built at several places along the river bank.

Even more alarming is the fact that the alluvial soil in the Ganges basin is showing signs of salinity in the state of Haryana, alkalinity in western Uttar Pradesh, acidity in West Bengal and calcareousness in north Bihar. Irrigation of Ganges water adds 17 million tonnes of salt a year to the soil, according to CPCB. The Ganges basin receives 3,000 tonnes of pesticides a year and one-third of all the fertilizer used in India.

K.S. Jayaraman

Polio eradication

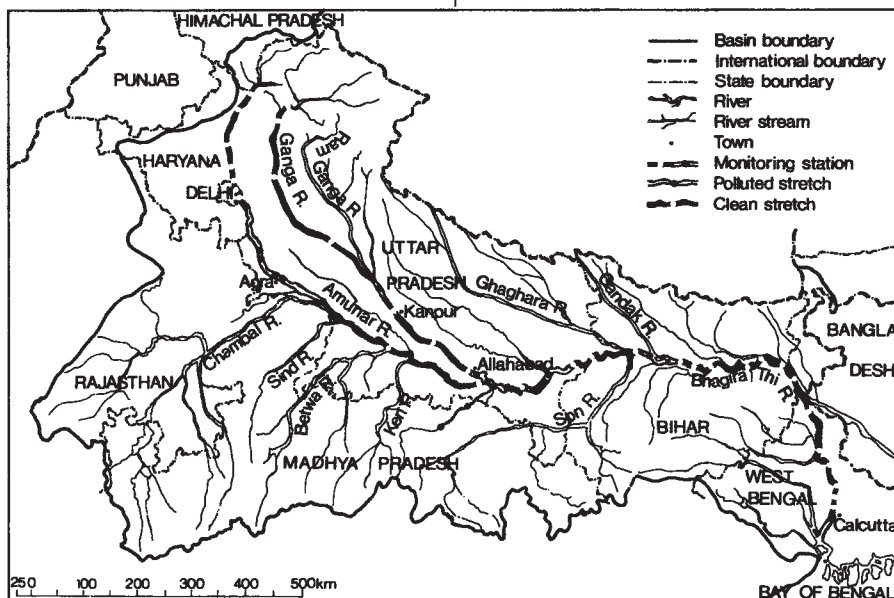
Washington

A PLAN to eradicate poliomyelitis from the Americas by the year 1990 has been announced by the Pan-American Health Organization (PAHO). PAHO believes the goal can be achieved by using its staff and volunteers to coordinate separate national efforts to encourage immunization.

The total cost of eradicating polio from North and South America is estimated at \$100 million: much of the infrastructure required is, however, already in place and only about \$30 million will be needed in aid from developed countries, much of which is already committed. Since 1977, when the World Health Organization instituted an expanded immunization programme to fight several common infectious diseases, the incidence of polio in the Americas has fallen from around 2,300 cases a year to 489 cases in 1984.

Dr Carlyle Guerra de Macedo, director of PAHO, says that 75 per cent of those under one year old in the Americas are now immunized against polio; before the expanded programme started, the figure was 25 per cent. The new plan will concentrate on surveillance and on training health workers to implement vaccination schemes. While admitting that the goal will not be easily achieved, Guerra de Macedo pointed out that the Americas were the first region of the World Health Organization to succeed in eliminating smallpox. The polio virus, like the smallpox virus, cannot survive for long periods outside a host and there is no animal reservoir. Total world elimination of polio is therefore a realistic goal.

Tim Beardsley



Running down observatories

SIR — As a university astronomer who has had much to do with both the Science and Engineering Council (SERC) observatories at all levels, including serving on committees charged with looking into whether they could best be merged with each other, or elsewhere, I would like to comment on your leading article on "Running telescopes" (*Nature* 9 May, p.86) which is a mine of misinformation.

Before there is any talk of closing or merging anything, with all the personal and professional upset and delay that will certainly entail, the case has to be made, and indisputably made, that there are significant inefficiencies in the present system of two observatories at Herstmonceux and Edinburgh, sharing some of their tasks, computing for example, with the Rutherford Appleton Laboratory. To that extent I agree with *Nature* that the widest terms of reference must be given to the new committee, if there has to be one so shortly after the last, which I doubt, looking into the best way of running our overseas telescopes.

Building and operating a high technology instrument on an exceedingly remote and sometimes physically hostile foreign site, transporting shifts of UK university astronomers there, looking after them, educating them in the intricacies of the local equipment, which has to be maintained at the highest efficiency, providing them with the technical help to get first rate data, and later the sophisticated computer network required to reduce such data, must be a technical and management challenge of the same order as running an aircraft carrier full of Harriers at war in the Falklands. Mistakes are inevitable, and any number of committees can be set up to look into how the task can be better done, and of course the less they know about the practicalities the easier it will be for them to recommend sweeping changes.

The professional and the user, however, will only be convinced by a proper comparison, carried out largely by observationalists, of the existing UK set-up with its competent counterparts abroad such as the Cerro Tololo Inter-American Observatory (US) and the European Southern Observatory based in Munich. My own survey, for what it was worth, showed the UK observatories in a very favourable light. If the new committee fails to make a similar but much more careful comparison, co-opting the advice of astronomers abroad, and taking the time to do it properly, they will certainly have failed astronomers, like myself, out in the UK universities. At present we are lucky enough to have access to shared facilities which are remarkable not for their size, as your leading article implies, but for their sheer quality, something in which both observatories, and the Rutherford Appleton Laboratory, can take a just pride. Such quality arises not from

lavish expenditure but from vision, high skill and dedication, as well as the very closest cooperation with the user community.

All this could be easily lost in the sort of reorganization or tidying up so beloved by (and alas in the United Kingdom indulged in by) second-rate administrators on their way up the civil service career ladder. One would have thought the amalgamations that led to British Leyland would have taught somebody up there a lesson.

The predictions are that the United States will be running short of astronomers in 2-3 years time. The demoralization that is now spreading through our UK observatories as a result of continual misinformed sniping by those who do not use them, and a never-ending series of superficial reviews, can in the end have only one result. And when the best SERC people have left on the plane, the rest of us, contemplating the resulting inevitable decline in the competitiveness of our own far from ambitious facilities (our colleagues abroad are going up to 10 or 15 metre class telescopes), will be studying the job columns too. To their credit, our observatories have played a major part recently in catapulting British optical and infrared astronomy from mediocrity into the front rank. In a society increasingly settling for the seedy and the second-rate perhaps that is their crime.

MICHAEL DISNEY

*Department of Applied Mathematics
and Astronomy,
University College,
PO Box 78,
Cardiff CF1 1XL, UK*

Philistinism?

SIR — If the idea of exploring more unorthodox ways of funding science than using the increasingly tired arguments for an ever-increasing Science Vote is the role of a scientific philistine then I must accept that label. What a pity though that a once great industrial society should display cultural antipathy to modern commerce of the sort revealed in your leading article (25 April, p.657).

My purpose in trying to provoke a debate on these important issues (not incidentally the unwitting act on my part that you seem to think) is to get scientists to consider a little more, in these financially difficult times, their role in society and their justification for that role being supported from the country's tax base. How you can contend that it is valid to challenge the scientific community with such arguments but not ask individual scientists to accept their implications is beyond me.

Of course I realize that not all science is commercially exploitable. Much of my own research on viruses has been in that category. But I still claim that there is a need for scientists to think purposefully about whether commercial opportunity ex-

ists and how it can be harnessed. Though there are some of us who need no such urging, this is not an area in which we have demonstrated much skill in the past. We might all benefit if more of us thought it important and generated income for ourselves, our institutions and our country as a result.

A major point in my *Times* letter that you did not choose even to misquote was that such a strategy could well generate funds for basic science both directly and from government. It is surely worth giving our thinking a nudge in this direction, to counterbalance the traditional arguments so frequently put, even at the risk of incurring the wrath of the Editor of *Nature*.

K.A. HARRAP

*Wolfson College,
Oxford OX2 6UD, UK*

Imunovir

SIR — Helen Wright's reply¹ fails to destroy Dimitri Viza's complaint² that the drug isoprinosine was hyped by excessive promotion, but rather confirms his implicit contention that the publicity for a product is proportional to the doubts about its activity. This contrasts with the case of an effective treatment for herpes; there has been no excessive publicity in the promotions of acyclovir, a drug whose virtues and limitations are well documented.

All but one of the references cited by Wright to prove the efficacy of Imunovir are to communications at meetings between 1973 and 1981, suggesting that none was followed up by a formal paper. She omits references to work suggesting the absence of effect^{3,4}.

There is a striking difference between Wright's arguments in defence of Imunovir — a 60 per cent cure of herpes bouts within a week, the praise of Prix Galien's jury for a drug arousing "new hopes for the treatment of infectious diseases, auto-immune diseases, allergy and — why not? — cancer" and the absence as yet of any convincing clinical trial. This is even more baffling if one considers that isoprinosine has been tested for more than 12 years.

I cannot therefore but agree that pharmaceutical companies would be well advised to avoid raising excessive hopes through lay media or otherwise. Since mediaeval times, people have been wary of panaceas. This is only one of the reasons why I am prepared to wager that in another five years isoprinosine will have fallen into oblivion. But Viza's proposal that there should be more scrutiny and responsibility in lay publications deserves to succeed.

JEAN-YVES FOLLEZOU

*Centre des Tumeurs,
Hôpital de la Pitié,
75013 Paris, France*

1. Wright, H.J. *Nature* 314, 126 (1985).

2. Viza, D. *Nature* 313, 344 (1985).

3. Kalimo, K.O.K., Joronen, I.A. & Havu, V.K. *Archs Derm.* 119, 463 (1983).

4. Glasgow, L.A. & Galasso, G.J. *J. infect. Dis.* 126, 162 (1972).

Piecemeal ballistic missile defence

The US programme of research known as the Strategic Defense Initiative will come to fruition only in the next century (if then), but there may be tangible and awkward results much sooner.

EVER since President Reagan's announcement in March 1983 of the plan for a programme of research on ballistic missile defence, there has been a temptation to assume that the whole programme would come to fruition (or not) all at once. Then, we have assumed, it would fall to a Pentagon committee to recommend (or otherwise) the deployment of an integrated system of anti-missile defence. And then it would be for politicians to take up the question which they have promised must be explored before such a defensive system could be put in place, negotiations with the Soviet Union not least among them. Only then, the argument has gone, would it be decided what should be done with the products of many years of research.

A little thought will show that events cannot follow such a simple course. The objective of the system now being talked about in Washington, in its most flamboyant form, is first to detect hostile ballistic missiles early in their trajectory and then to destroy them by means of several layers of defence. If boost-phase missiles escape attack, there would ideally be means of destroying the components into which they separate, first "buses" containing several warheads and, later, individual re-entry vehicles. Much of the argument of those working on the programme rests on the calculation that the most effective route to a missile defence will be an efficient way of destroying missiles while their booster motors are still burning, and are thus conspicuous targets for the detector satellites equipped with infrared sensors.

To say the least, each successful strike against a hostile missile in this boost stage would get rid of a great many warheads. But it is also generally agreed that the detection problem is also then much easier than during the intermediate stages (when decoys will be an obvious stratagem for the missile launchers to use). Indeed, detection becomes relatively simple again only towards the end of a trajectory, when warheads declare themselves by the heat they generate.

No secret knowledge is required to make reasonable guesses about the sequence in which the various components of this system will become available to the military. The technology for detecting missiles during the boost phase probably requires nothing radically new. Infrared detectors have indeed been used to spot rocket boosters rising through the atmosphere. There is no reason why detection

should not be reasonably efficient at heights above the ground not very different from those at which people operate infrared telescopes, a few kilometres or so.

The practical task will be to incorporate an array of such detectors into a system of satellites and then handle the data that accumulate. Discrimination between real and imagined targets will obviously be a headache. Software designers will be fully stretched. But the first practical by-product of the research programme will be an early warning system of such sophistication that it might have been considered desirable in its own right.

The other almost certain outcome of the research programme is a better system of terminal defence than that planned by the United States in the late 1960s. The idea then was that a fast take-off missile (called "Sprint") would fire nuclear warheads at incoming warheads; now, it seems, people are persuaded that it would be better to send guided mechanical devices at incoming warheads, in the hope that they might be destroyed without killing the people whose defence is the object of the whole exercise.

So the research programme will almost certainly yield two military options for the United States — a much better (in the sense of much earlier) early warning system and a workable system of terminal defence. Whether the research programme will contribute anything else is still an open question.

The important article in last week's *Nature* (23 May, p.286) by Dr Richard Garwin is a generalized expression of the kind of scepticism with which the research programme will have to contend. Any system of satellites intended to be the bases ("battle stations" is the fashionable term) for the means by which targets will be destroyed will have to be much more elaborate than the enthusiasts have been saying. Even quite optimistic assumptions about the efficacy of lasers as means of shooting down hostile missiles imply much larger numbers of battle stations than people have been saying.

The need for a host of satellite battle stations is generally accepted. Because these machines, whatever their armoury (lasers, neutral-beam projectors or rail guns), will have to be in comparatively low orbits to get within striking distance (say 1,000 km) of a booster rising through the atmosphere, the chances are that any single one of them will be the wrong side of the Earth when

needed. But the ballistic missile defence community has sought some comfort in the widely shared belief that the number of satellites required will be proportional only to the square root of the number of potential targets. Garwin disproves that cheerful assumption.

This is the technical nightmare of the whole project. Even though the mass-production cost of making, say, 1,000 battle stations will yield economies of scale, the efficient management of such a system would be a task of quite unprecedented character. Only a means of destroying boosters at much greater range than 1,000 km would bring drastic simplification. Because none of the potential guns has yet emerged as a possible leader, let alone been demonstrated, it seems a fair guess that this part of the programme will be much delayed. No Pentagon committee is going to be making a recommendation to deploy the system this side of the end of the decade. Whether that time will come this century is far from certain.

The consequences of this piecemeal fruition of the various components of the programme are themselves important. If, as seems likely, one early outcome is a feasible system of terminal defence, it seems likely that the United States will follow the Soviet Union in deploying such a system somewhere, presumably in the Minuteman missile fields where the MX missile is now to be deployed. (The Anti-Ballistic Missile Treaty originally offered each side the option of two such systems, later reduced to one.) An orbiting early-warning system, the other likely early product of the research programme, will be diplomatically more perplexing. Quite apart from the formidable technical difficulties of launching, operating and maintaining such a system, there seems very little doubt that it would violate the spirit if not the letter of the treaty.

The result is that, while the Strategic Defense Initiative will not yield tangible results for a decade or more ("if then", as the critics would add), there is every likelihood that some essential parts of the system will be ready for use long before that. So the United States may quite soon find itself having to keep to the promise extracted from the administration by Mrs Margaret Thatcher last December that deployment would depend on talks with the Soviet Union. Raising the questions now, at the arms control talks at Geneva, would be prudent.

John Maddox

Plant genetics

A plant joins the pantheon at last?

from Geoffrey North

As Keith Roberts has recently remarked in *Nature*¹, plants should be wonderful systems for the study of development. Why, then, has progress in understanding cell differentiation and morphogenesis in plants been so slow? One reason, surely, has been the lack of a plant equivalent of the fruitfly *Drosophila melanogaster*: a rapidly breeding species that is easy to keep in large numbers in the laboratory, and is ideal for intensive study by the combination of molecular and genetic techniques that has proved so successful in dissecting the regulation of development in *Drosophila*². But it seems plant scientists need look no further, for the indications from a recent meeting* are that the small weed *Arabidopsis thaliana* (the common wall cress; see sketch) could prove to be just the right subject. And the hope that molecular biology will transform practical plant breeding must have been boosted by reports at the same meeting of the correctly regulated expression of exogenous genes introduced into plants using the natural vector system provided by the bacterium *Agrobacterium tumefaciens*.

Although the science of genetics began with Mendel's experiments with peas, on the whole plants are far from ideal subjects for molecular geneticists. They usually require a lot of space and special conditions for their upkeep, and, worse, as a rule undergo only one or at most a few breeding cycles in a year. Furthermore, the species on which plant geneticists have concentrated most of their efforts, such as maize, tend to have enormous genomes — in the region of two orders of magnitude larger than those of *Drosophila* or yeast, which is a considerable handicap if your aim is to isolate a specific gene. These considerations led Elliot Meyerowitz (California Institute of Technology) to the temperate weed *Arabidopsis thaliana*, a member of the Cruciferae (mustards). Although the first *Arabidopsis* mutations were reported in the 1940s and it has since been used extensively in the genetic analysis of plant biochemistry (C. Somerville, Michigan State University; see ref. 3 for a recent genetic map of *Arabidopsis*), its potential as a subject for intensive molecular genetic analysis seems hitherto to have largely been unexploited.

The features that attracted Meyerowitz to *Arabidopsis* are the ease with which it can be grown in the laboratory, its small size and extraordinarily short generation time (only 4–5 weeks), and its unusually small genome⁴. On the basis of DNA reassociation kinetics Meyerowitz estimates that the *Arabidopsis* genome is only ~70,000 kilobases (kb)⁵, the smallest yet

for any plant and comparable to the yeast, nematode and *Drosophila* genomes. For comparison, the size of the human genome is ~2,000,000 kb, and that of rye is ~7,900,000 kb. Furthermore, the DNA reassociation kinetics show that the *Arabidopsis* genome is remarkably low in repetitive sequences, which should facilitate the application of the 'chromosome-walking' technique for isolating genes that made possible the cloning of the bithorax complex of *Drosophila*. Meyerowitz hopes that the small genome of *Arabidopsis* will rapidly be covered with sites of restriction-fragment length polymorphisms — his group have already detected a few by simply using randomly selected clones from an *Arabidopsis* genome library. These can be used to map new mutations and as starting points of chromosome-walks to clone the sites of the mutations.

The economy of the *Arabidopsis* genome is also reflected at the level of specific genes; for example, where most plants have multiple copies of the genes for seed proteins, *Arabidopsis* has, according to Meyerowitz, only one. And using their cloned *Arabidopsis* seed protein gene as a probe, Meyerowitz's group have carried out the first analysis of gene expression in

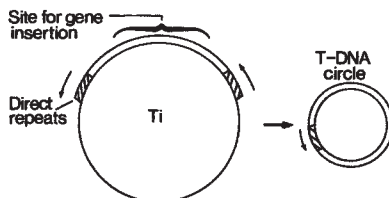
plants by hybridization *in situ*. In sections of seeds, they can detect seed protein messenger RNA only in embryo cells, which are known to be the sites of seed protein synthesis in *Arabidopsis*. The result in itself is to be expected; what is important is that it shows the powerful technique of hybridization *in situ* can be applied in *Arabidopsis*. It is not always easy to make it work and this has, for example, been the principal barrier to progress in research on homoeotic gene expression in vertebrates.

Arabidopsis seems to be as amenable to the techniques of cell culture, plant regeneration and *Agrobacterium tumefaciens*-mediated transformation⁶ as plants such as tobacco (see below and box). One of the problems that has frustrated plant scientists in recent years has been the lack of any one species that is both amenable to such techniques and has been well-studied genetically. And although I. Potrykus (Friedrich-Miescher Institute) and E. Howard (CSIRO) both reported that they have been able directly to introduce DNA into protoplasts of monocotyledonous plants, the difficulties in regenerating such plants from cells in culture make it likely that *Arabidopsis* will be the first genetically well-defined plant into which exogenous genes can be introduced readily.

The most exciting prospect offered by *Arabidopsis* is in its potential for the identification, genetic mapping and cloning of those genes which, on the basis of the effects of their mutations, seem likely to be very closely involved in the regulation of

Plant transformation

To date, the success of plant geneticists in introducing new genes into plants has depended on the exploitation of a remarkable natural system of interspecific gene transfer. The bacterium *Agrobacterium tumefaciens* can induce the formation of tumours on many plant species. This is known to involve the transfer of a fragment of DNA (the T-DNA), carried on the bacterium's tumour-inducing (Ti) plasmid, to the genome of the infected plant cell. It is the genes carried by the T-DNA that cause the transformed cells to divide to form tumours and that pervert the metabolism of the tumour cells so that they make products on which the bacterium can feed, such as nopalines.



As recently reported in *Nature*¹, circular molecules of T-DNA are induced within the bacterial cells in response to some plant-derived signal. It is presumed that these molecules represent intermediates in the process of gene transfer — though as yet no T-DNA circles have

been detected in plant cells. The junction of the T-DNA circles occurs at a precise point within the short (25 base pair) direct repeats that flank the T-DNA on the Ti plasmid, and simply by inserting a piece of DNA between these repeats it can be transferred to the genome of a plant cell. So that whole plants can be regenerated from the transformed cells, 'oncogenic' genes of the T-DNA are usually removed and replaced by a selectable marker, such as an antibiotic-resistance gene.

Unfortunately, it has so far not been possible to use this system to transform monocotyledonous plants, despite the indications that they can be infected by *Agrobacterium tumefaciens*^{2,3}. One alternative approach that is being pursued at present is to see whether transposable elements could be used to introduce genes into plants, much as P-elements are used to transform *Drosophila*. As a start in this direction, R.B. Simpson (ARCO, Dublin, California) reported at the plant genetics symposium that his group has begun to explore the potential of the maize *Mu* (Robertson's *mutator*) transposable elements as vehicles for transforming monocots.

Geoffrey North

*The 1985 UCLA symposium on plant genetics was held at Keystone, Colorado, USA from April 13–19 1985.

1. Koukolikova-Nicola, Z. *et al.* *Nature* 313, 191 (1985).
2. Hooykaas-Van Slogteren, G.M.S., Hooykaas, P.J.J. & Schilperoord, R.A. *311*, 763 (1984).
3. Hernaesteens, J.-P., Thia-Toon, L., Schell, J. & Van Montagu, M. *EMBO J* 3, 3039 (1984).



ARABIDOPSIS
THALIANA

Redrawn from *The New Concise British Flora* by W. Keble Martin (Ebury Press and Michael Joseph, London, 1982).

development. Several such mutations were described in Colorado, for example, by D. Meinke (Oklahoma State University), who explained that he and his colleagues have identified a number of embryonic-lethal mutations that cause arrest at different stages of development by exposing *Arabidopsis* plants to the mutagen EMS and then screening the seed pods for aborted seeds. Meyerowitz's group have reanalysed a mutation termed *apetala-1*, classically described as causing a failure of petal development; they find that occasionally petals are replaced by entire new buds.

The *apetala-1* mutation is somewhat reminiscent of the homeotic mutations of *Drosophila*, in which one body part may be transformed into another. Perhaps more closely comparable, at least at the phenotypic level, are the maize mutations *Corngrass* and *Teopod 1* and *2*, described by S. Poethig (University of Pennsylvania).

The repetitive structure of a maize shoot, with units (termed phytomers) comprising a leaf, a node and a portion of stem, is at least superficially similar to segmentation in animals, the pattern of which can be affected by homeotic mutations. In maize, at fairly precise positions up the plant, the character of the phytomers changes: from vegetative at the bottom through female flower (ear)-forming to male flower (tassel)-forming at the top. The effects of *Corngrass* and *Teopod* mutations can be interpreted as transforming upper phytomers into lower phytomers, as shown by leaf-shape and the replacement of flower parts by vegetative structures.

Another maize gene worth mentioning in this context is the *R*-locus, which is not only intriguing in its own right but illustrates another strategy for the cloning of plant genes. The interest in the *R*-locus lies in the observation that mutations in different regions of the locus specifically affect the coloration of different parts of the plant so, as J. Kermicle (University of Wisconsin) explained, it seems likely that the *R*-locus is somehow involved in regulating the tissue-specific pattern of expression of the enzymes of anthocyanin synthesis in maize.

S. Dellaporta (Cold Spring Harbor Laboratory) reported that the insertion of a transposable *Ac* (*Activator*) element into the *R*-locus of mutant maize has allowed the cloning of the locus — the *Ac* element had already been cloned, so could be used as a stepping-stone to clone the adjacent *R* DNA.

If the *R*-locus really is a 'master' regulatory gene, then its most likely targets would be DNA sequences that determine the pattern of expression of the genes for the enzymes of anthocyanin synthesis. The technique of introducing defined fragments of exogenous DNA into plants by *Agrobacterium tumefaciens*-mediated transformation, alluded to briefly above, offers a direct route to the identification of such sequences, and the power of the approach was well illustrated at the meeting. R. Goldberg (UCLA) reported that when fragments of DNA including either the trypsin inhibitor or seed lectin genes of soya bean are introduced into tobacco cells and whole plants are regenerated from the transformed cells, the exogenous genes are developmentally regulated just as they are in soya-bean plants. RNA transcripts can be detected in developing embryos at the correct times during embryogenesis, and in other tissues either no transcripts or the same low levels that occur in the equivalent tissues of soya bean are found. And when a large fragment of soya-bean genomic DNA, containing the seed-lectin gene and several neighbouring genes with different development profiles of expression, is incorporated into tobacco plants each gene is differentially regulated in the trans-

formed plants in precisely the same way as in the soya bean. Similarly, Sengupta-Gopalan *et al.*⁷ have found that when the gene for the major seed storage glycoprotein (β -phaseolin) of French bean is introduced into tobacco plants, on a fragment of DNA that includes the β -phaseolin coding-region with about a kilobase of 5' flanking sequence, it retains its normal pattern of expression: β -phaseolin can first be detected in the seeds of the transformed plants at the time it normally would appear in developing French bean seeds, several days after the tobacco plants' own storage proteins have started to accumulate.

The implication of these results is that the fragments of soya-bean and French-bean genomes introduced into the tobacco plants include the information to specify in which tissue the genes should be expressed. The definition of the sequences that comprise this information should be relatively straightforward; but to know how to direct the expression of any gene to seed-tissue should be an exciting prospect for those who wish to use molecular genetic techniques to improve crop plants. □

1. Roberts, K. *Nature News and Views* 314, 581 (1985).
2. North, G. *Nature News and Views* 311, 214 (1984).
3. Koornneef, M. *et al. J. Hered.* 74, 265 (1983).
4. Bennett, M.D. & Smith, J.B. *Proc. R. Soc. B274*, 227 (1976).
5. Leutwiler, L.S., Hough-Evans, B.R. & Meyerowitz, E.M. *Molec. gen. Genet.* 194, 14 (1984).
6. Aerts, M., Jacobs, M., Hernalsteens, J.-P., Van Montagu M. & Schell, J. *Plant Sci. Lett.* 17, 43 (1979).
7. Sengupta-Gopalan, C., Reichert, N.A., Barker, R.F. Hall, T.C. & Kemp, J.D. *Proc. natn. Acad. Sci. U.S.A.* (in the press).

Geoffrey North is Deputy Biology Editor of *Nature*.

Planetary science

More news from Mars

from I.P. Wright and M.M. Grady

SAMPLES of Mars have been delivered to the Earth as meteorites, but this has been appreciated only in the past few years, following, among other things, the discovery of meteorites from Antarctica that have a demonstrable lunar origin. The stimulus this has provided for the discipline of Solar System geochemistry was clearly evident at a recent conference on lunar and planetary science*. Four shergottites (S), three nakhlites (N) and one chassignite (C) together make up the eight so-called SNC meteorites thought to originate from Mars. Detailed studies of these, and in particular the results of a consortium investigation of the Shergotty meteorite (organized by J.C. Laul, Battelle Pacific North West Laboratory, Washington), have thrown new light on martian history as well as the processes by which the SNC meteorites themselves were delivered to Earth. The SNC meteorites provide a bridge between laboratory-based analyses aimed at elucidating the origins and evolution of the

Earth, Moon and meteorites, and planetary geology.

Research groups at the Battelle Pacific North West Laboratory, Max-Planck-Institut für Chemie, Mainz and Purdue Universities (headed by J.C. Laul, H. Wänke and M.E. Lipschutz, respectively) have made detailed chemical analyses of the Shergotty meteorite, which provide constraints on the bulk composition of the parent body. Their deductions of the composition of the parent body comply with that inferred from geophysical data (density, moment-of-inertia factors, etc.) concerning the core/mantle structure of Mars. There are subtle variations in chemical composition that demonstrate that Shergotty is heterogeneous on a centimetre scale; for example, there are large variations in concentrations of volatile elements such as cadmium, tellurium, thallium and bismuth. The separated minerals contain widely differing proportions of rare-earth elements (REE), with phosphates such as apatite and whitlockite containing 500 times the proportions found in chondrites. In addition,

*The 16th lunar and planetary science conference was held in Houston, Texas, 11-15 March, 1985.

the light REE seem to be concentrated in the iron-rich rims of clinopyroxene crystals. It is clear from the trace-element data that Shergotty and EETA 79001 (an Antarctic shergottite) are not related to the same parent magma. Further chemical studies show a high abundance of tungsten in Shergotty (which has a molybdenum/tungsten ratio of 1.0, compared with 5.9 for the terrestrial mantle); this seems to accord with various processes thought to operate during core formation and is consistent with the view that Mars was formed by homogeneous accretion.

Previous oxygen isotope measurements of whole-rock samples have demonstrated that SNC meteorites form a distinct group that probably originated on a single parent body distinct from the Earth-Moon system. The oxygen isotope composition of mineral separates of Shergotty has been determined by R.N. Clayton and T.K. Mayeda (University of Chicago). Data from mineral pairs give equilibration temperatures of $1,100 \pm 100^\circ\text{C}$; the degree of isotope fractionation between the mineral pairs is lower than that often seen in equivalent terrestrial materials and demonstrates that little, if any, secondary exchange with water has taken place following crystallization. Not all the minerals analysed conform to this interpretation. Indeed, a maskelynite (shocked plagioclase) separate from Shergotty has an extremely anomalous oxygen isotope composition, although this cannot be due to the shock process itself because maskelynite from other shergottites shows no unusual isotope composition. The 0.4 per cent enrichment of ^{18}O in the maskelynite from Shergotty remains an intriguing and unexplained result, possibly reflecting low-temperature alteration.

Nitrogen in shocked glass from EETA 79001 has previously been shown to have an isotope composition enriched in ^{15}N , consistent with an origin from the martian atmosphere. A portion of the nitrogen in Shergotty also appears to be of atmospheric origin. Two groups, R.H. Becker and R.O. Pepin (University of Minnesota) and R.H. Carr, I.P. Wright and C.T. Pillinger (Open University), have found evidence for a nitrogen component in Shergotty that shows a 6 per cent enrichment of ^{15}N . It is thus considered that the shergottites probably contain microscopic and heterogeneously distributed shock-produced glasses that in turn contain samples of the martian atmosphere. J. Yang and S. Epstein (California Institute of Technology) and Carr *et al.* find that some of the carbon in Shergotty is enriched in ^{13}C , consistent with both isotope measurements of CO_2 in the shock-produced glass in EETA 79001 and by *in situ* measurements of the martian atmosphere by the Viking spacecraft. Thus, some of the CO_2 in Shergotty also seems to be of martian atmospheric origin, although the picture is somewhat confused by the presence of a number of indigenous carbon-bearing components that need to be further resolved before the carbon

chemistry of the SNC meteorites is completely understood.

The glass fraction of EETA 79001 has been subjected to detailed optical examination by D.W. Muenow and J.L. Gooding (University of Hawaii; NASA Johnson Space Center, Houston). Using scanning electron microscopy, they have recently identified two distinct components in the glass fraction — a smooth glass and a 'ribbed' glass, which is thought to be quench-textured augite. Vesicles have also been identified in the glasses, as have sulphur-rich aluminosilicate crystals that are finely disseminated throughout the samples. It is possible that these secondary minerals are products of Antarctic weathering, although they may have originated on another volatile-rich planet, such as Mars.

J.H. Jones (University of Arizona) stirred up controversy by proposing that the previously accepted chronology of the shergottites should be called into question. Briefly, the current dogma is that the shergottites have an igneous origin at 1,300 Myr, then underwent impact metamorphism and subsequent ejection from the martian surface at 180 Myr. The new hypothesis, based on a reappraisal of existing trace element and isotope data, suggests that the shergottites were actually formed

in an igneous event about 180 Myr ago. Furthermore, the cosmic-ray exposure age of shergottites (2.5 Myr), previously thought by most people to represent the time the meteorite fragmented in space, may actually approximate to the time the samples were ejected from Mars. The shorter exposure age of EETA 79001 would, however, require two separate ejection events for the shergottite meteorites, which poses problems as it demands high cratering rates on Mars over the past few million years.

The results obtained from studies of SNC meteorites have widespread implications for theories of Solar System formation, planetary accretion, thermal histories of planets and petrogenesis on planetary bodies. Chemical and isotope analyses of SNC meteorites have given a clearer appreciation of the composition of the martian surface and atmosphere. There is a strong feeling in the community of lunar and planetary investigators that a Mars sample-return mission should be given high priority in space exploration programmes in the years ahead. □

I.P. Wright and M.M. Grady are in the Department of Earth Sciences, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK.

Developmental neurobiology

What guides growing axons?

from Dennis Summerbell and R. Victoria Stirling

DURING normal development, neurones form processes that grow long distances to connect with their targets. They produce a precise and repeatable pattern that is essential for normal functioning of the nervous system. There are numerous theories about the mechanisms involved in forming this pattern and three papers in this issue illustrate different ways of looking at the development of these connections, in tissue culture, in explants and *in vitro*. None of the papers present new explanations but all are important because they involve observing events as they happen, rather than at a fixed end point.

By removing nerve cells into tissue culture, one can control most carefully for critical single factors. We already know that axons prefer to grow along particular substrates, for example, laminin, but can one demonstrate that they have specific preferences? On page 409, Bonhoeffer and Huf describe a Y-maze in which embryonic chick retinal axons from one edge of the eye are given the choice of growing along axons from the same position or axons from the opposite edge 180° away. Axons from one particular location (temporal retina) clearly prefer to grow along their neighbours, whereas axons from the opposite (nasal) side show no such preference. Here, in a dish, we can watch axons exhibiting the kind of behaviour suggested

from static observations of their trajectories in fixed amphibian brains, in which temporal axons are found in a tight bundle in the optic tract while those from nasal retina, 180° away, are distributed across its entire width. The development of the highly ordered projection from the eye to the brain may depend on such differences, using them, for example, to set up the orientation of the map, and this kind of experiment has the potential for analysing the molecular basis for the differences.

But such fundamental yes/no decisions are only part of the story; a second approach is systematically to remove putative guidance cues and observe the resulting pattern of axon growth. The developing insect wing provides a culture system for the study of nerve growth that is less abstracted from the whole animal. The wing primordium starts life as a hollow flattened disk (the imaginal disk), rather like a pitta bread, in which the guidance of axons from peripheral sensory cells can be followed. Blair, Murray and Palka (on page 406) have adopted this model to look at the role of the substrate in axonal navigation. Axons from the sensory cells in the wing follow stereotyped pathways towards the central nervous system. Among the possible cues they might use are the channels between the dorsal and ventral surfaces of the wing, but when the disk is split into

separate dorsal and ventral halves, the dorsal axons are still able to find their way despite the absence of channels to guide them. Another possible cue is the nearby neurones, which in whole embryos appear to act as 'stepping stones' or 'guideposts'. Laser destruction of these guideposts leads to disoriented growth of specific axons (Raper, J.A., Bastiani, M.J. and Goodman, C.S. *J. Neurosci.* 4, 2329; 1984). In the present experiments, the removal of the proximal guidepost neurone did not prevent the axons of more distal cells from moving initially in the right direction.

In both these experimental systems, molecular dissection could be contemplated, but one must remember that cells in culture are not in their natural environment. Some of the important cues may not be expressed, whereas others may not be relevant *in vivo*. The finer intricacies of connections depend on the interaction of axons with their continuously changing environment, so that the ultimate level of observation must be the intact, living animal. Recent advances in techniques for intracellular dye injection coupled with video-image enhancement have been harnessed brilliantly by Purves and Hadley on page 404 of this issue. The superior cervical ganglion in the neck of the rat is an accessible cluster of cells nicely suited to this technique; by using image intensification, the cell can be photographed without obviously harming it. They first expose the ganglion, fill a particular cell with fluorescent dye and photograph its branching pattern. Several days or even a month later they are able to find the same cell again and repeat the performance. They observe that the branches of the neurone change their shapes progressively over time: nerve cells are not static but are continually adapting, even in the adult animal.

It is precisely in these types of developmental systems that molecular techniques need to be applied. Increasingly, we will identify the appearance of cell surface and extracellular matrix constituents by the use of monoclonal antibodies and discover unique, regionally-localized transcription by *in situ* hybridization. Such results must be correlated with the kind of developmentally interesting events described in these papers. For example, the appearance of cell-adhesion molecules provides a powerful tool to study development both at the level of cells in the organism and at the level of regulation of gene expression in the cell (Edelman, G.M. *Proc. natn. Acad. Sci. U.S.A.* 81, 1460; 1984).

It is becoming increasingly apparent that no single factor is responsible for the development of specific neural connections. In the intact animal, there seems to be a redundancy of guidance cues. These papers show that the intact organism must not be forgotten in the quest for molecular answers to developmental questions. □

Dennis Summerbell and R. Victoria Stirling are at the National Institute for Medical Research, Mill Hill, London NW7 1AA, UK.

Vision

Disparities in depth perception

from C. Longuet-Higgins

ONE might be surprised that 150 years after the invention of the stereoscope anything remains to be discovered about stereopsis, but on page 402 of this issue Graeme Mitchison and Suzanne McKee describe a novel effect called 'interpolation', which they observe when looking at a particular kind of stereogram. In the past few years stereopsis has had an injection of new interest from two developments in particular: the demonstration by B. Julesz and J. Chang. (*Biol. Cybernet.* 33, 107; 1976) that pattern perception is not a prerequisite for stereoscopic fusion; and the so-called 'computational paradigm' — the recognition that vision is the outcome of an elaborate neuronal computation for which our eyes supply only the input.

The foremost exponent of the computational paradigm was the late David Marr, whose book, *Vision* (Freeman, San Francisco, 1982), dealt in some detail with stereopsis. With Tomaso Poggio he produced, in succession, two alternative theories for the mechanism of stereopsis. Though the details of the process are still largely obscure, few visual psychologists would deny Marr's premise that an important preliminary to stereoscopic fusion is the matching of corresponding elements in the two retinal images — the well known 'correspondence problem' of stereopsis. Once the elements have been matched, the visual system can assess the disparities between their retinal positions and draw conclusions about the relative distances of the visible features of the scene.

The observations of Mitchison and McKee bear most directly on the question of how the visual system carries out the matching process. Most of their stereograms, presented for a time too short for the observer to scan the display, were interpreted exactly as predicted by standard stereoscopic theory. The apparent difference in depth between a pair of dots in the stereogram was simply proportional to their horizontal disparity, that is, the difference between their horizontal distances apart in the left and right retinal images. This presupposes that the elements of the two images have been successfully matched; otherwise their disparities are ill-defined.

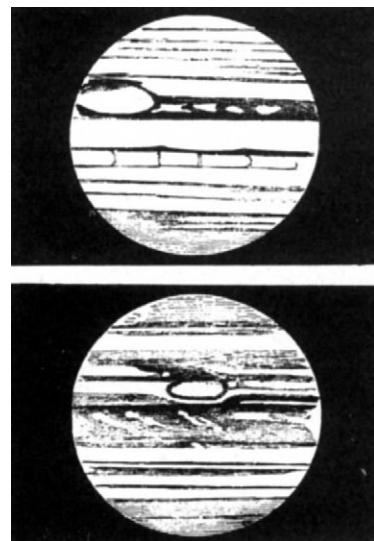
Erratum

Due to an editorial mistake in the article by A.R. Gardner-Medwin 'Potential challenge from glia' (*Nature* 16 May, page 181), it was implied that several long-established properties of glial cells were new findings reported in recent papers. The lines "including . . . fascination" should be removed from the penultimate sentence of the first paragraph; instead they should complete paragraph 2, which will then read ". . . biological cells, high permeability . . .".

With the stereograms that gave predictable impressions of depth, the only uncertainty in the outcome arose from alternative ways of matching the images; the choice between such alternatives seems to be broadly consonant with what is known about the 'disparity gradients' associated with matched pairs of visible elements (Burt, P. and Julesz, B. *Science* 208, 615; 1980). But the stereograms that Mitchison and McKee find most interesting are those containing regular grids of very closely spaced points, the critical spacing being about 6' of arc — roughly the angle subtended by adjacent grooves of a nail file held in the hand. At this spacing the visual system seems to 'lose count' of the total number of points in each image, and to estimate the position of the grid by the horizontal disparity between its two edges, ignoring any differences in the numbers of intervening points in the two images; the whole grid is seen as roughly coplanar.

Mitchison and McKee point out that their observations cast doubt on Marr and Poggio's concept of a 'coarse-to-fine vergence strategy', in which the correspondence problem is solved by progressively finer adjustments in the directions of gaze of the two eyes. It will be interesting to see whether they can develop a computational theory of the matching process with a predictive power equal to that of the theories they challenge. □

C. Longuet-Higgins is in the Laboratory of Experimental Psychology, University of Sussex, Brighton BN1 9QG, UK.



Probably recurrent markings on Jupiter. Top, 27 November 1857 (Dawes); bottom, 12 h 50 min on 25 February 1885 (Denning). From *Nature* 32 33, 14 May 1885.

Geophysics

Electrical earthquake prediction

from Paul W. Burton

Two thousand years ago, Aristotle proposed that earthquakes were caused by subterranean winds. The Greeks are still active in the often contentious business of earthquake prediction and the team of P. Varotsos, K. Alexopoulos and K. Nomicos are well aware of controversy surrounding their work. In two papers which appeared recently in *Tectonophysics*^{1,2}, Varotsos and Alexopoulos have put forward their observations and ideas on variations in the Earth's electric field preceding earthquakes. They now claim to be able to forecast the magnitude and epicentre of earthquakes in Greece; to prove this they have taken the unusual step of issuing and publishing registered telegrams of their predictions, to establish a check on their work.

Phenomenological earthquake prediction is essentially the study of precursory phenomena — those events that precede an earthquake. These range from peculiarities in animal behaviour and lights in the sky to the more geophysically obvious and comprehensible precursors of increased minor seismicity. The work of Varotsos and Alexopoulos is confined to electro-telluric currents in the Earth. These naturally occurring electrical currents, which are not fully understood, are measured by probing the surface at two or more points, approximately 30 metres or more apart. They might be 'channelled' along conductors within rocks, such as fluid-filled cracks; because earthquakes involve changes in the stresses in the ground and ultimately movement along fractures, it is plausible that changes in the electro-telluric regime could indicate imminent earthquakes.

This team has monitored the telluric field at several sites in Greece since March 1981 and searched for precursory seismic electric signals (SES). Their most important conclusion is "that every sizeable earthquake is preceded by an SES and inversely every SES is always followed by an earthquake, the magnitude and the epicentre of which can be reliably predicted". These seismic electric signals are observed 6–115 hours before an earthquake, although there are distinct groupings within this range. Varotsos and Alexopoulos find no correlation between either the SES lead time or duration and the magnitude of the ensuing earthquake. However, SES from several earthquakes at different distances from a station show changes in electrical potential that are inversely proportional to distance; the SES amplitude from one seismic region recorded at one station shows a simple log-linear relationship to magnitude. Comparison of information from several stations enables calibration of earthquake signals and prediction.

Specialists and non-specialists will rake

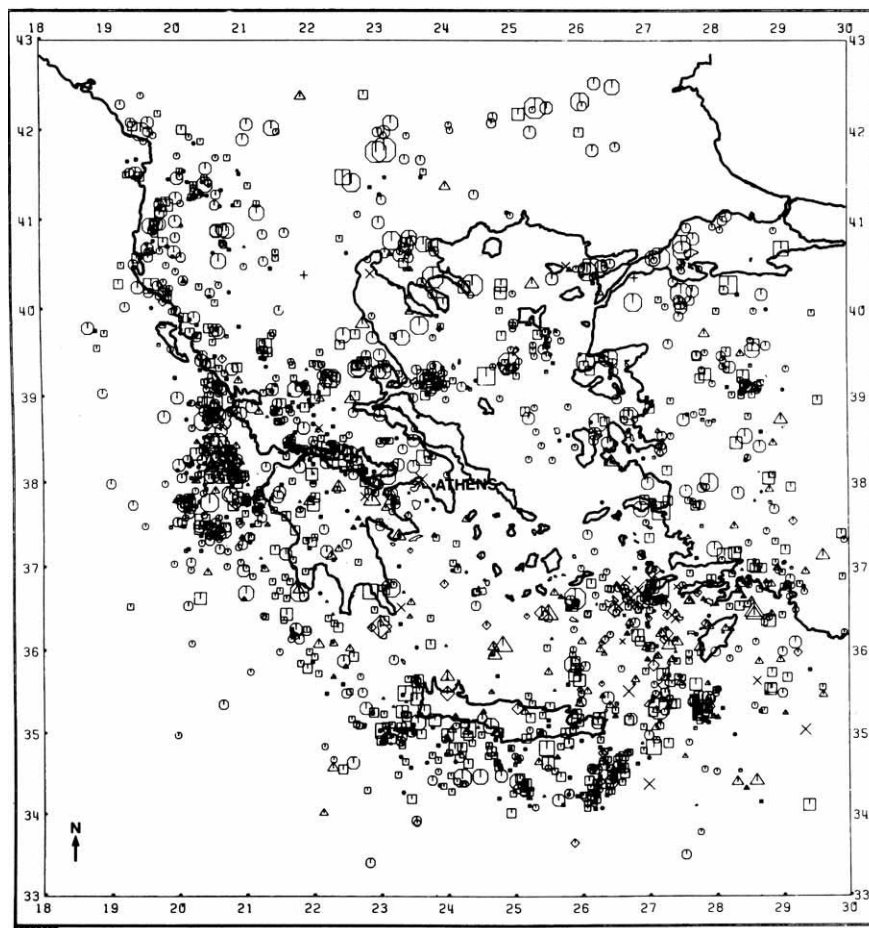
through these papers to judge how far the work is advanced and how justified are the publicized conclusions. The reproduced charts showing the SES variation with time are often confusing at first sight and do not present their information clearly. But they are reminiscent of an era when seismograms themselves looked primitive to the uninitiated.

Calibration of one station in western Greece involved collecting 20 SES measurements over four months from earthquakes exceeding magnitude 4 in the Kefalonia (Cephalonia) region and comparing them with magnitudes determined by the National Observatory of Athens. A few hundred earthquakes above this magnitude were probably recorded within 100 kilometres of Cephalonia during this time and one wonders how Varotsos and Alexopoulos dealt with all these other SESs, which they presumably also recorded. Consequently, it seems odd that there is little comment on the overall seismicity of the region. Greece is the most seismically active area in Europe (see figure), and there are clear dangers in invoking correlations

between frequently occurring events. When Varotsos and Alexopoulos put forward their case, backed by registered telegrams, for having forecast 21 out of 23 earthquakes of magnitude 5 or larger that occurred near their network during eight months in 1983, they do not emphasize that one of these was forecast with magnitude 4. Taking this lower threshold, there were probably a few hundred earthquakes within a radial distance of about 1.5 degrees of the station at Pirgos in the western Peloponnese in the same period. With so many earthquakes, many will seek further verification of the correlation between particular earthquake, SES and telegram, and it would also be interesting to know how many telegrams contained information which was found subsequently to be incorrect or fortuitously correlated.

Nevertheless, many people are hopeful that a significant breakthrough has been made. This prompts consideration of the general problems raised by earthquake prediction. There will doubtless be engineering seismologists, earthquake engineers, land-use planners and government decision-makers who would hope that attention and funds are not diverted from other more practical aspects such as continued programmes of house strengthening or better building codes in earthquake-prone zones.

There is also a degree of *a posteriori* demonstration involved, which raises the



Distribution of seismicity in Greece (1901–1977). Lines in symbols are longer for greater magnitude of disturbance; symbols range from circles (surface earthquakes) to crosses (deepest earthquakes).

very real consideration of what to do when a prediction concerns an earthquake that might have substantial consequences. The Seismological Society of America has issued guidelines³ which state that "a scientist has fulfilled his or her obligations by informing colleagues, employer and the appropriate public agencies of the prediction hypothesis, its basis, and any critiques, concurrences or nonconcurrences based on independent reviews". The United States also has a formal mechanism for dealing with earthquake predictions, in which the responsibility for informing both the public and public officials of vetted earthquake predictions is vested in the Director of the US Geological Survey. Community preparedness for earthquakes would also seem to be a prerequisite. In Japan, the Earthquake Assessment Committee is backed up by the Large-Scale Earthquake Countermeasures Act⁴. Greece has recently followed these leads and established, by Act of Parliament, the Organization for Aseismic Protection. But

when the lead time for an electrotelluric prediction is only 6–115 hours and it may relate to any part of Greece, it is hard to see what might be expected of such an organization.

So, inevitably, what should be done next? There is obviously a case for continued investment in seismological studies and earthquake source physics, but such investigations alone do not allow people to sleep well in their beds at night. In the long term, the less glamorous role of the earthquake engineer and the careful implementation of building codes and environmental hazard maps may perhaps be more substantial, if unsung, life-savers. □

1. Varotsos, P. & Alexopoulos, K. *Tectonophysics* 110, 73 (1984).

2. Varotsos, P. & Alexopoulos, K. *Tectonophysics* 110, 99 (1984).

3. Guidelines for Earthquake Predictors in *Bull. seism. Soc. Am.* 73, 1955 (1983).

4. Earthquake Prediction and Preparedness in Japan in *Earthquake Inf. Bull.* (U.S. Dep Int. Geol. Sur.) 15, 142 (1983).

Paul W. Burton is at the British Geological Survey, West Mains Road, Edinburgh EH9 3LA, UK.

Neurochemistry

New light on synapsin I

from Richard Rodnight

MANY crucial processes in the nervous system, such as synaptic transmission, involve complex interactions between the neuronal plasma membrane and components of the neuronal cytosol. In the search for the molecular mechanisms of such processes, the identification and characterization of nerve-specific proteins is an essential step. There have been few successes in this area, but one that stands out is the discovery over a decade ago by P. Greengard and colleagues¹ of a major phosphorylated protein in brain, originally termed protein I, but now called synapsin I. Subsequent work by Greengard's group elegantly demonstrated that synapsin I occurs exclusively in the nerve terminals, where it is bound to the surface membranes of the synaptic vesicles, the storage sites for neurotransmitters². This highly specific location clearly pointed to an important role for the molecule in some aspect of synaptic transmission (for example, the still enigmatic mechanism for transmitter release) but no precise evidence for its function emerged. Now, A.J. Baines and V. Bennett, on page 410 of this issue³, present evidence to show that synapsin I is very similar, if not identical, to an important spectrin-binding cytoskeletal protein termed brain 4.1.

The significance of this exciting discovery may reside in the fact that the analogous protein in erythrocytes, erythroid 4.1, can form stable ternary complexes with spectrin and the contractile protein actin. In the erythrocyte, these complexes constitute a stable supportive lattice at the inner face of the cell membrane and are

therefore important in maintaining the shape of the cell; in nerve terminals (where spectrin and actin also occur) they may play a more dynamic role in modulating the interaction between synaptic vesicles and the terminal membrane.

To appreciate the significance of this development, we must take a closer look at both the cytoskeletal proteins and synapsin I. In the neurone a key molecule is brain spectrin⁴ (also called fodrin or calspectrin), a tetramer formed from two subunits of approximate relative molecular mass (M_r) 260,000 and 265,000, antigenically related to erythrocyte spectrin (for further comment on spectrins as a family of proteins, see ref. 5). Similar to its red-cell analogue, brain spectrin binds both actin and protein 4.1 (or synapsin) and is in turn anchored through its smaller subunit to the inner surface of the neuronal membrane by yet another cytoskeletal protein, ankyrin⁴. Synapsin I consists of two subunits of approximate M_r 80,000 and 85,000 and, as noted by Baines and Bennett³, has been mainly studied as a substrate for cyclic AMP-dependent and calcium/calmodulin-dependent protein kinases⁶. Each subunit comprises a globular head and a highly basic tail with phosphorylation sites on both domains. The basic tail of the molecule is bound to the external surface of the synaptic vesicle membrane by ionic forces that are weaker when the tail is phosphorylated by the calcium/calmodulin-dependent protein kinase¹.

It is particularly interesting that changes in the phosphorylation state of synapsin induced by electrical stimulation of nerves

or by drugs that stimulate by depolarizing the neuronal membrane seem to involve mainly the calcium/calmodulin-dependent phosphorylation. This makes it possible that a consequence of such stimulation (which may be the prelude to neurotransmitter release) is to break the bond between synapsin I and the vesicle membrane. Previously, the significance of this reversible binding of synapsin I to vesicles was obscure, but now that Baines and Bennett have shown that the molecule also binds to the neuronal cytoskeleton through spectrin, intriguing possibilities open up. For example, by analogy with the interaction of protein 4.1 and spectrin in erythrocytes, the binding of vesicle-associated synapsin I to brain spectrin may greatly enhance the capacity of brain spectrin to crosslink actin filaments, thereby inducing a change in the shape of the sub-membranous cytoskeleton and re-orientating the vesicle in relation to the terminal membrane. Further changes in vesicle orientation might follow disruption of the synapsin I – vesicle bond as a consequence of the calcium/calmodulin-dependent phosphorylation of synapsin I induced by calcium-ion fluxes in the nerve terminal. It may be significant that spectrin has binding sites for calmodulin.

Many questions need to be answered however, before precise roles for the synapsin-spectrin interaction in synaptic transmission can be formulated. First, it should be pointed out that Bennett's group have so far been unable to demonstrate an effect of synapsin I on the crosslinking of actin by spectrin, although other workers have reported such an effect with erythrocyte protein 4.1 (ref. 7). Perhaps the explanation of the discrepancy lies in differences in the function of the spectrin-actin-4.1 complex in the red cell and nerve terminal — in the former the need for a stable supporting structure is evident, whereas in the latter a more dynamic function may be envisaged, possibly involving transient states that would be difficult to detect in the test tube. Second, the precise intraneuronal distribution of brain spectrin is far from clear. As Baines and Bennett point out, spectrin is axonally transported and spectrin-like molecules can associate with intracellular organelles⁸. Moreover, the two subunits of spectrin do not always occur together and the larger subunit has a wider distribution⁹, although both are needed to crosslink actin filaments. Nevertheless, there is enough evidence to suggest that spectrin, perhaps together with tubulin, could form part of a much more extensive cytoskeletal network than a purely membrane-associated structure. Such a network might serve to regulate the disposition and movement of synaptic vesicles in the nerve terminals by binding vesicle-associated synapsin I.

A final conundrum concerns the fact that both spectrin-like molecules and proteins immunologically related to erythrocyte 4.1 have a wide distribution in cells; indeed

Goodman *et al.*¹⁰ present evidence indicating that, at the level of the light microscope, the distribution of spectrin and 4.1 (as detected by an antibody to erythrocyte 4.1) in the cerebellum are identical and extend to neuroglia as well as neurones. The latter contrasts with the virtually exclusive localization of synapsin I in the nerve terminal¹. These results cannot be explained purely on the basis of the very high concentration of synapsin I in the nerve terminal (where it constitutes about 6 per cent of the vesicle protein), because immunocytochemistry using anti-synapsin immunoglobulin confirms the restricted location. Therefore, it seems likely that synapsin I is a highly specialized form of 4.1 found only in neurones. This conclu-

sion confirms the overwhelming impression that synapsin I must play a unique role in some aspect of synaptic transmission. □

1. Huttner, W.B., Schiebler, W., Greengard, P. & De Camilli, P. *J. Cell Biol.* **96**, 1374 (1983).
2. Navone, F., Greengard, P. & De Camilli, P. *Science* **226**, 1209 (1984).
3. Baines, A.J. & Bennett, V. *Nature* **315**, 410 (1985).
4. Bennett, V., Davis, J. & Fowler, W.E. *Nature* **299**, 126 (1982).
5. Baines, A.J. *Nature News and Views* **312**, 310 (1984).
6. Nestler, E. & Greengard, P. *Nature* **303**, 583 (1983).
7. Burns, N.R., Ohanian, V. & Gratzer, W.B. *FEBS Lett.* **153**, 165 (1983).
8. Aunis, D. & Perrin, D. *J. Neurochem.* **42**, 1558 (1984).
9. Cumming, R. & Burgoyne, R.D. *Biosci. Rep.* **3**, 997 (1983).
10. Goodman, S.R., Casoria, L.A., Coleman, D.B. & Zagon, I.S. *Science* **224**, 1433 (1984).

Richard Rodnight is in the Department of Biochemistry, Institute of Psychiatry, De Crespigny Park, London, SE5 8AF, UK.

Solid-state physics

How superconductivity fades near metal-insulator transitions

from W.L. McLean

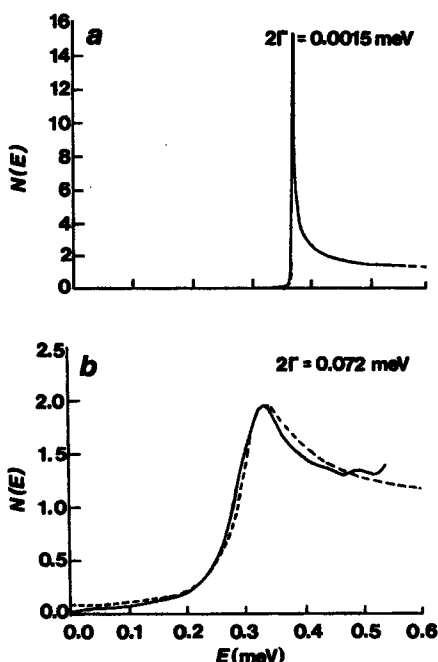
ALTHOUGH the importance of quantum mechanics in solid-state physics has been well established for many decades, it came as a rude surprise about five years ago that startling quantum effects in metallic conduction had been overlooked. Unusual behaviour and bizarre predictions near the metal-insulator transition have recently come to be expected. Further challenging questions arise when the system can also be superconducting. One of these that is less innocent than it first seems is whether superconductivity can persist into the insulating side of the transition. If not, does it die out right at the transition or while the system is clearly still on the metallic side? R.C. Dynes, J.P. Garno, G.B. Hertel and T.P. Orlando have recently shed new light on this area by measuring electron tunnelling in the granular aluminium system (*Phys. Rev. Lett.* **53**, 2437; 1984).

Granular aluminium, which is a mixture of grains of metallic aluminium and amorphous aluminium oxide, is one of a number of systems from which a series of specimens, ranging from pure metallic to pure insulating, can be prepared by varying the metal volume fraction. The granularity, which can be seen by transmission electron microscopy, is clearly marked in the middle range of concentrations, where the metal-insulator transition occurs. Many theoretical models that attempt to explain properties of granular metals do explicitly take this granularity into account.

A very different interpretation that has been used by Dynes *et al.*, among others, is that the properties of granular aluminium can be interpreted as if the system were an amorphous metal in which the disorder is on an atomic scale. This outlook is commonly justified by the assumption that all relevant length scales are larger than the grain size, and it is in this framework that

the importance of the new tunnelling measurements becomes apparent.

The experiments use the technique of electron tunnelling at low temperatures of approximately 0.06 K, which is much less than the superconducting transition temperature, to measure the density of states of the elementary excitations (quasi-particles) of the superconductor. In low resistivity specimens it is not surprising that



The density of states, $N(E)$ deconvoluted from the experimental data. The dashed line is a BCS density of states broadened by the value of Γ shown in each figure. The 'lifetime' of a quasi-particle is $\hbar/\Gamma$. *a*, Specimen with small broadened effects; *b*, specimen very close to metal-insulator transition (modified from Dynes, R.C. *et al. Phys. Rev. Lett.* **53**, 2437; 1984).

granular aluminium behaves exactly as an ordinary superconductor with a sharp asymmetrical peak in the density of states clearly defining the edge of an energy gap, in good agreement with the Bardeen, Cooper, Schrieffer (BCS) theory of superconductivity. At higher resistivities very close to the metal-insulator transition, however, the density of states is no longer zero at low excitation energies and the discontinuity is replaced by a broad smooth maximum.

Dynes *et al.* have successfully fitted their data with a BCS density of states modified to take into account broadening caused by a finite lifetime of the quasi-particles in the granular superconductor. In this case the lifetime is limited by inelastic scattering processes, which have been of great interest recently in connection with the tendency of electron states to become localized as the metal-insulator transition is approached. The values of the inelastic scattering rates deduced from the tunnelling measurements are consistent with those deduced from a different experimental approach, measurement of the magnetoresistance.

Thus, superconductivity seems to die away near the metal-insulator transition by a gradual lifetime broadening of the density of states concomitant with the appearance of quasi-particle states in the region where there is an energy-gap in lower resistivity specimens. A particular oddity of these results is that they imply that moderately strong indications of superconductivity still appear close to the energy at which the sharp peak occurred in the low resistivity specimens. This may be connected with the long tails reported by Dynes *et al.* in the resistive transitions of the highest resistivity specimens.

Further intriguing questions are: first, what happens at much lower temperatures when the lifetimes of the quasi-particles should become longer? Does the full BCS gap reappear? The implication is that the closer one approaches the metal-insulator transition, the lower must the temperature be in order to see the full superconductivity. Second, does some inelastic scattering mechanism peculiar to the granular structure come into play or are the mechanisms found in amorphous metals still operating in the granular system?

Although the new results seem to provide firm support for the quasi-homogeneous medium approach, it still remains to reconcile this point of view with approaches that do assume the granularity to be relevant. These include those workers who seek to explain the properties of granular metals in terms of percolation through an assembly of randomly linked grains, and those focusing on the consequences of electrostatic charging of the grains on the stability of the superconductivity. □

W.L. McLean is visiting the Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, UK.

Explanation of the inward displacement of Io's hot plasma torus and consequences for sputtering sources

J. A. Linker,* M. G. Kivelson,* M. A. Moreno* & R. J. Walker

Institute of Geophysics and Planetary Physics and * Department of Earth and Space Sciences, University of California, Los Angeles, California 90024, USA

Radial profiles of the ion density and flux-tube content in the Io torus have peak values inside Io's orbit, even though Io is the effective source of these ions. Formation of an inward peak constrains either the velocity distributions or source regions of sputtered neutrals. A further constraint is that the ionization of neutrals on trapped trajectories that return to Io's surface must be limited. A dominant sulphur source is most easily reconciled with these constraints.

Io, the innermost of the four Galilean satellites, has a dense torus of plasma associated with it. This torus, perhaps the most intriguing feature of Jupiter's magnetosphere, has been the subject of intensive research and debate. Recent ground-based measurements of the torus¹, as well as measurements from the Voyager spacecraft^{2,3}, have shown that both the ion density (Fig. 1a, b) and NL^2 , the flux-tube content (Fig. 1c), attain their maximum values as functions of radial distance from Jupiter well inward of Io's orbit, with NL^2 falling off sharply inside ~ 5.6 jovian radii (R_J).

Many features of the radial profiles can be understood in terms of radial diffusive transport⁴⁻⁷. Because diffusion driven by macroscopic processes in planetary magnetospheres is accomplished by the interchange of magnetic flux tubes, NL^2 , proportional to the number of ions per unit magnetic flux, is the diffusing quantity⁴. (The parameter L labels a field line and, in a dipole field, L is the distance in planetary radii from the centre of the planet to the equatorial point on the field line. N is the number of ions per unit L .) The radial profile of NL^2 is consistent with diffusive transport of ions only if their source attains a maximum value near 5.6 – $5.7 R_J$. Centrifugal forces drive rapid outward diffusion, consistent with the relatively slow outward decrease of NL^2 beyond $5.7 R_J$. Slow inward diffusion driven by atmospheric winds explains the sharp decline inward from $5.6 R_J$, but the heavy ions that form the torus must originate from matter on Io itself. The main purpose of this paper is to explain why, although Io is the actual source of particles, the peak in NL^2 occurs inside its orbit. Specifically, we have calculated the trajectories of neutral particles ejected from Io and have examined the conditions under which they tend to be ionized inside Io's orbit. We find that the requirement of a localized peak constrains either the spectrum of sputtered neutrals or their source locations on Io, or both. A further constraint is placed on the ionization of neutrals on trapped (non-escape) trajectories.

Calculations and modelling

It is generally agreed that the ions in the torus originally are ejected from Io as neutrals forming neutral clouds extending far from Io. These neutrals are ionized eventually to form the torus⁸⁻¹⁰. This sequence suggests a possible explanation for the peak in NL^2 inwards of Io's orbit; perhaps the trajectories of neutrals are such that they tend to be inward of Io's orbit at the time they are ionized¹. We have investigated this possibility using simple models for the production of ions from the neutrals. We have assumed that neutral particles are ejected from Io's equator with initial positions spaced uniformly in longitude and separated by 5° (see Fig. 2a). At each initial position, 100 neutral particles are 'launched' radially, with specified initial velocities,

each set of particles having a distribution of lifetimes of the form (Fig. 2b)

$$P(t) dt = (dt/\tau) \exp(-t/\tau) \quad (1)$$

We ignore the consequences of misalignment of magnetic and rotational poles. Thus, the position of a neutral at the end of its life defines the radial coordinate of a newly created ion. To be physically meaningful, it is important that a distribution of lifetimes be used for the model, because if only the average lifetime were used there would be no way for the particles from a given source position with given velocity to spread in radial-position.

At each initial position, trajectories were calculated for velocities in the range 1.0 – 4.0 km s^{-1} in increments of 0.1 km s^{-1} . We used an orbit calculation program (OBCAL) that integrates the equations of motion for the particles, taking into account the gravitational forces of Jupiter and Io, the other three Galilean satellites and the J_2 term in Jupiter's gravitational potential. (The Galilean satellite trajectories are modelled as slightly inclined ellipses.) Ephemeris data for the Galilean satellites were provided by the Jet Propulsion Laboratory (JPL). The program was tested by comparing calculations of the positions of both Io and the Galileo spacecraft with calculations performed by JPL; discrepancies were $< 5 \times 10^{-3} R_J$.

Sputtering, in which charged particles collide with neutral matter and neutral atoms and molecules are emitted, is believed to be the dominant process responsible for ejecting matter from Io¹⁰. To calculate the number of particles at a given velocity, we have assumed velocity distributions expected for sputtering. The particles are assumed to come from the surface of Io; however, this should not be construed as eliminating the possibility that the particles could originate from the atmosphere or exobase. As any assumed surface source distribution maps to a shifted velocity distribution at the atmosphere or exobase, our solutions also apply to the mapped sources at these altitudes. Note that all our sources are located at Io's equator and Io is assumed to lie in the magnetic equatorial plane; these assumptions serve to make the calculation manageable. High-latitude particles emitted radially would have similar trajectories except that their orbits would be inclined with respect to the spin equator. Because energy is used to incline the trajectory rather than to change the apsides of the trajectory, high-latitude particles in general need higher initial velocities to escape and to achieve the same eccentricity as equatorial particles. Therefore, the trajectories of high-latitude particles should be similar to those of equatorial particles, except that the orbits will be inclined and higher initial velocities will be necessary to achieve the same displacement from Io's orbit. By neglecting the tilt of Jupiter's dipole, we do not take into account that the L shell of

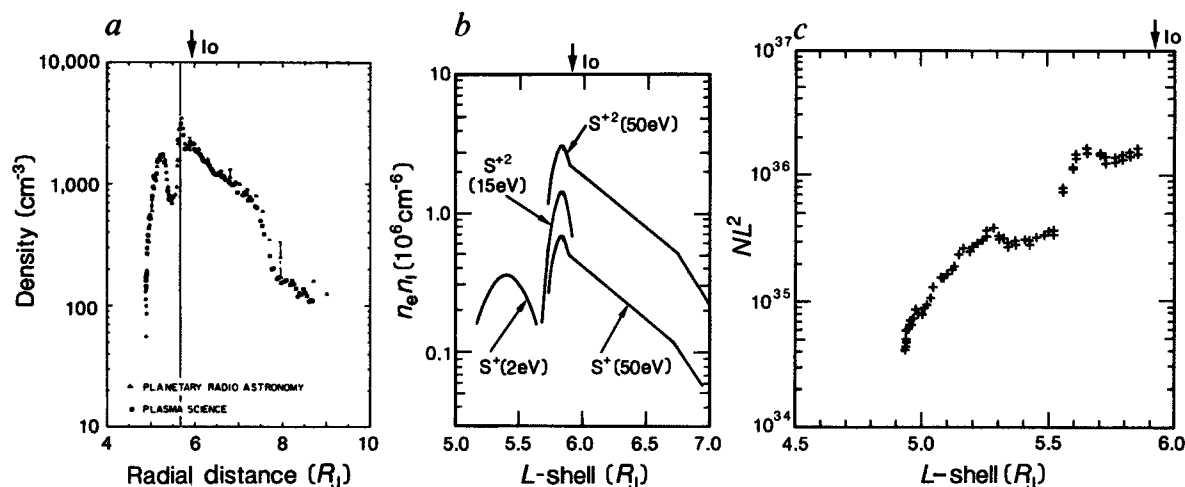


Fig. 1 *a*, Radial profile of *in situ* measurements from the Voyager spacecraft of charge density in the Io torus. The peak occurs well inside Io's orbit, at $\sim 5.7 R_J$ (adapted from ref. 3). *b*, Radial profiles of ground-based measurements of sulphur-ion densities. Again the peak occurs inside of Io's orbit (adapted from ref. 1). *c*, Radial profile of NL^2 , proportional to the total number of ions per unit magnetic flux, based on measurements from the Voyager spacecraft. This function, conserved in adiabatic transport, enters the canonical form of the equation for radial diffusion. The peak in NL^2 occurs at $5.6\text{--}5.7 R_J$ and thus requires a source of ions in this region. There also seems to be a 'plateau' at $5.2\text{--}5.3 R_J$ in the cold torus (this also appears as a small peak in *a* and *b*; adapted from ref. 2).

a newly created ion is necessarily the same or larger than the radial distance at which ionization occurs. However, because the electron density (and hence the ionization rate) is much higher in the centrifugal equator¹¹, neutrals will tend to be ionized when they are in the centrifugal equator. The change in L shell is then very slight ($<0.02 R_J$) and should not have any significant effect on the profiles shown here.

Results

Unless otherwise noted, our calculations assume a spatially uniform probability of ionization. First, we consider a mean lifetime of 40 h, appropriate for sulphur atoms being ionized by electron impact near Io. This is the lifetime that results if we use commonly accepted rate coefficients¹²⁻¹⁴ and an electron density of $\sim 2,000 \text{ cm}^{-3}$. Figure 3*a* shows the radial distribution of new ions from neutral particle sources distributed from 0 to 90° in longitude. Neutral velocities are assumed to obey what is called a thermal spike velocity distribution¹⁵

$$f(v) dv \propto v^3 \exp(-mv^2/kT) dv \quad (2)$$

and only neutrals above the escape velocity ($v > 2.3 \text{ km s}^{-1}$) are considered. (Here T , an arbitrary parameter characterizing the steepness of the distribution and not related directly to the actual temperature of the source, was chosen as $4,000 \text{ K}$. The thermal spike velocity distribution is only an approximation to observed sputtering spectra. For example, values of $2,000$, $3,600$ and $5,200 \text{ K}$ have been used to represent spectra for sputtering of various materials¹⁶.) Note that a peak in density occurs between 5.6 and $5.7 R_J$ as required by the data. One should also recognize that the plotted profiles in Figs 3-5 are proportional to the source rate in a diffusion equation for NL^2 .

Figure 3*a* shows that sputtered neutrals above the escape velocity with a plausible velocity distribution produce the spatial localization needed for a source of Io torus ions. The neutrals that tend to contribute most to the ionization peak between 5.6 and $5.7 R_J$ are those with a velocity slightly above the escape velocity ($2.3\text{--}2.5 \text{ km s}^{-1}$). This can be understood as follows: particles that barely escape from Io, once away from Io, are on nearly circular trajectories around Jupiter and thus tend to be ionized at roughly the same distance independent of the lifetime of the neutral. Those neutrals that are injected inward thus tend to be ionized inwards of Io's orbit in a fairly narrow range of radial position. For higher injection velocities, neutrals tend to be on more elliptical trajectories and are thus ionized with a

wider spread in radial position. As the thermal spike velocity distribution heavily favours injection of neutrals at lower velocities, neutrals with injection velocities favourable for formation of the inward peak tend to dominate the escaping particles. The $0\text{--}90^\circ$ region of longitude is also favourable for formation of the desired peak because neutrals ejected from this region proceed on trajectories inwards of Io's orbit. However, as will be shown later, a source of neutral particles spread uniformly in longitude can also form the required inward peak. Our parameters are chosen to describe sulphur; however, because the results are relatively insensitive to lifetime they probably apply to oxygen as well.

Note that when all important gravitational forces are considered, particles with velocities as low as 2.3 km s^{-1} can escape Io even though the escape velocity from Io as an isolated body is almost 2.6 km s^{-1} . The number of neutrals ejected above the escape velocity gives the only constraint on the T parameter in the thermal spike velocity distribution. With $T = 4,000 \text{ K}$, as in this calculation, $\sim 4\%$ of the neutrals escape. As will be shown below, this gives an adequate supply rate of neutrals.

Thus far, we have considered only neutral particles on trajectories that allow them to escape from Io. However, there is nothing in our model to prevent the ionization of neutrals on trapped trajectories (particles that, if not ionized, ultimately fall back to the surface of Io). Because the sputtering velocity distributions favour low velocities, a large fraction of neutrals will be on these non-escape trajectories. If we consider the same velocity distribution as in Fig. 3*a* with the same longitudinal region of sources, but include the ionization of neutral particles below the escape velocity, the radial distribution of newly created ions shown in Fig. 3*b* is formed. A small peak forms between 5.6 and $5.7 R_J$, but contrary to the observations, a sharp peak forms at $5.9 R_J$. This sharp peak forms because many more neutral particles are on trapped trajectories than on escape trajectories. Because of Jupiter's gravitational force, these low-velocity particles get much farther away from Io and take much longer to fall back to Io's surface than if only Io's gravitational force were considered. We are presently working on calculations of the neutral density near Io and find that the requirement of an adequate source rate of neutrals above the escape velocity implies that a dense neutral cloud ($\sim 10^5\text{--}10^6 \text{ cm}^{-3}$) forms within 3 Io radii of the surface of Io. For the case shown in Fig. 3*b*, a source rate of $3 \times 10^{28} \text{ ions s}^{-1}$ into the torus implies that $>10^{28}$ ionizations s^{-1} occur near Io, greater by an order of magnitude

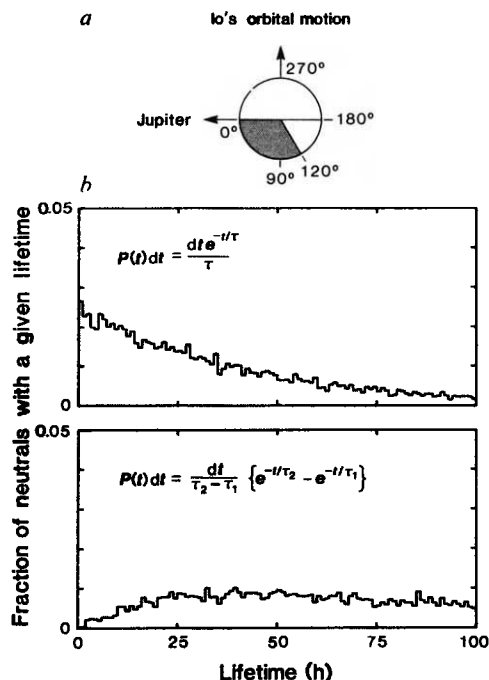


Fig. 2 *a*, Coordinate system for neutrals ejected from Io. The circle represents Io seen from above Io's orbit plane. For the calculations, particles were ejected uniformly in longitude with a separation of 5° between each position. The shaded area of the diagram shows the region where sulphur volcanoes are observed on Io²⁵. Note that the coordinate system is right-handed, opposite to the one used by ref. 25. *b*, Histogram of a typical distribution of lifetimes, generated for a mean lifetime against ionization of 40 h for a spatially uniform probability of ionization. The distribution gives the probability that an atom survives a time t and is ionized between t and $t+dt$. The radial position of a neutral at the end of its life defines the radial coordinate of the newly created ion. *c*, As *b* for the case of a molecule that is first dissociated and the products subsequently ionized: τ_1 , mean lifetime for dissociation of S_2 , was assumed to be 46 h (see text); τ_2 , the mean lifetime for ionization of S , is 40 h. A spatially uniform probability of electron impact is assumed. As the probability of ionization in the first few hours is small, the number of ions formed from neutrals on trapped trajectories is relatively insignificant. Because any S_2^+ ions formed are lost quickly by dissociative recombination, only S_2 molecules that are first dissociated contribute significant numbers of ions to the torus.

than the limit placed on ionization of sulphur and oxygen atoms in the vicinity of Io by observation¹⁷.

As the Io torus requires an injection rate of $2-7 \times 10^{28}$ ions s^{-1} (refs 18–20) and the sputtering velocity distributions greatly favour the production of particles on trapped trajectories, some process must limit the ionization of atoms on these non-escape trajectories. The ionization limit is required both to explain the formation of an inward peak and to be consistent with the observations of limited ionization near Io. One way that the ionization of atoms near Io could be limited is if the predominant sputtering products are molecules. If an electron impact is required to first dissociate the molecule, and the products are ionized by subsequent electron impacts, then neutrals on trapped trajectories would fall back to the surface of Io before they could become ionized.

For estimates of the importance of this two-step process we consider the sputtering of sulphur. Sulphur sputtered from the surface of Io has often been ignored as a source of torus particles because some estimates^{10,15} indicated that sputtering yields for this element would be small compared with those for SO_2 . These sputtering yield estimates were based on the assumption that the binding energy is much higher for sulphur than for SO_2 (ref. 10). However, recent experiments²¹ have shown that sulphur gives anomalously high sputtering yields of the order of 10^4 – 10^6 per incident ion when bombarded with 30–300 keV He^+ ions, a

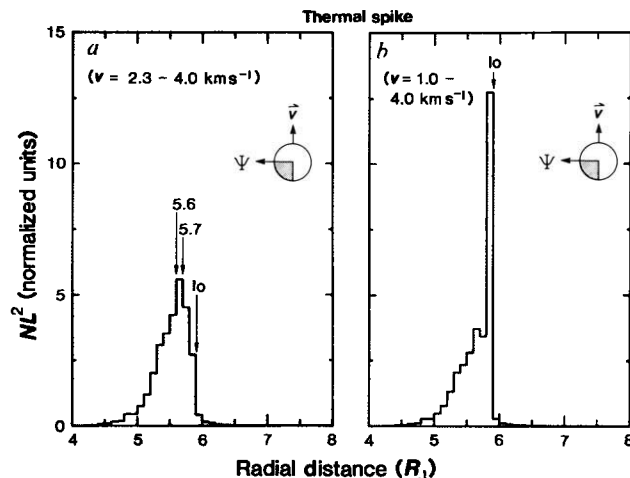
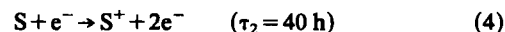
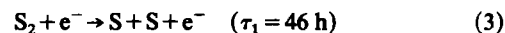
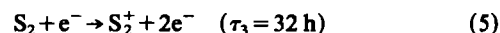


Fig. 3 *a*, Radial profile of NL^2 divided by the total number of particles, calculated for neutrals ejected from longitudes between 0 and 90° on the surface of Io with a thermal spike distribution of velocities, considering only velocities above the escape velocity. A distribution of lifetimes as shown in Fig. 2*b* with a mean lifetime of 40 h was used for the neutrals. The peak at 5.6 – $5.7 R_J$ is the major feature required of a source of Io torus ions. For neutrals above the escape velocity, any velocity distribution of ejected neutrals that falls off rapidly enough with increasing velocity will cause the inward peak in NL^2 to form. A rapid fall off with increasing velocity is in fact a characteristic of velocity distributions expected for sputtering. This figure and all succeeding figures show possible source profiles of Io torus ions. The profile of the total ion density would be modified by the effects of diffusive transport, but the peak would remain in the same location. *b*, As *a* except that neutrals with velocities below the escape velocity are included. A sharp peak forms at $5.9 R_J$, in contrast with the observations. If we normalize the number of neutrals above the escape velocity to be $2 \times 10^{28} s^{-1}$, an adequate source rate²⁰ of 3×10^{28} ionizations s^{-1} in the torus is obtained, but the profile shown here would require 10^{28} ionizations s^{-1} within 3 Io radii of Io's surface. Observations place a limit of 10^{27} ionizations s^{-1} of sulphur and oxygen atoms near Io. Thus, for consistency with observations, some process must limit ionization of sulphur and oxygen atoms on trapped trajectories near Io. Injection of S_2 neutrals as the dominant source of torus ions is consistent with this requirement. The profiles shown can be regarded as the source rate in a diffusion equation governing NL^2 .

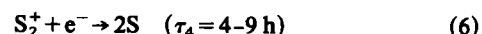
yield as much as five orders of magnitude greater than estimates for the sputtering yield of SO_2 (ref. 10). The experiments indicate that the predominant product of sputtering of sulphur is S_2 , with some larger sulphur molecules present as well. S_2 would dissociate and ionize as follows (the mean lifetime, τ , is shown for each process); because lifetimes for S_2 ionization and dissociation were not available, lifetimes for O_2 were substituted²²



On the other hand, the S_2 molecule itself can be ionized



There is no observational limit on the ionization per second of S_2 , as the observations near Io constrained only the ionization rates of atomic sulphur and oxygen. However, experiments with O_2^+ and NO^+ show that dissociative recombination occurs rapidly for these molecules for electron temperatures in the range 0.045 – 4 eV (ref. 23). Extrapolation of the dissociative recombination rates to 5 eV gives a lifetime in the range 4 – 9 h (for an electron density of $2,000 \text{ cm}^{-3}$), which is small compared with the ion diffusion time (30 – 100 days). Thus, the reaction



proceeds sufficiently rapidly that any S_2^+ ions formed will be converted into neutrals. The neutrals will then have the corotation velocity and will quickly leave the torus. Therefore, S_2^+ ions

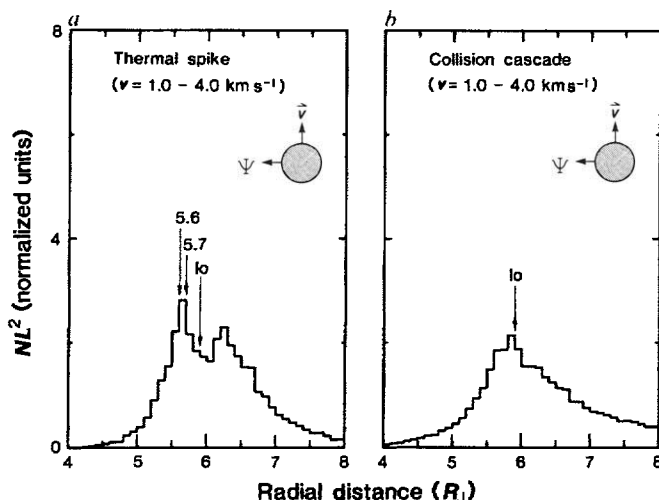


Fig. 4 *a*, Radial profile of NL^2 divided by the total number of particles. For this and all succeeding figures, the distribution of lifetimes described in Fig. 2c was used. Neutrals were ejected from all longitudes of I_0 , with a thermal spike distribution of velocities (see text). The range of velocities was $1.0-4.0 \text{ km s}^{-1}$. The required inward peak forms at $5.6-5.7 R_J$. This is true for any velocity distribution that peaks below the escape velocity and falls off rapidly with increasing velocity. *b*, As *a* but for the limiting case of a collision cascade distribution (see text). A peak at I_0 is formed that is inconsistent with the required source. This velocity distribution does not fall off rapidly enough to form the inward peak, but does provide a direct source of neutrals over a broad radial range.

will not add significantly to the NL^2 profiles, and we need consider only those neutrals that are first dissociated and then ionized.

For a spatially uniform probability of electron impact, the probability of a neutral's being ionized between time t and $t + dt$ following an earlier dissociation is of the form:

$$P(t) dt = \frac{dt}{\tau_2 - \tau_1} (e^{-t/\tau_2} - e^{-t/\tau_1}) \quad (7)$$

If we generate neutral lifetimes using this probability distribution (Fig. 2c) and we use the thermal spike velocity distribution that we used in Fig. 3, but with neutral sources at all I_0 longitudes, we obtain the radial distribution of newly created ions shown in Fig. 4a. As the dissociation of the molecules must occur before the individual atoms are ionized, only a fraction of particles on trapped trajectories are ionized and the peak at $5.9 R_J$ is reduced. A clear peak of ion density forms at $5.6-5.7 R_J$, as required by the data.

For the results shown in Fig. 4a, we have looked at a specific mechanism that can limit the ionization rate near I_0 if the source is rich in S_2 . However, other processes can limit the ionization rate near I_0 independently of the neutral species. Because the neutral cloud within 3 I_0 radii of I_0 's surface has neutral densities orders of magnitude higher than the electron density, the physics must differ greatly in the neutral cloud from elsewhere in the torus. Other, possibly more important, mechanisms may limit ionization rates for neutrals in the dense cloud, and thus decrease the ionization probability for particles on trapped trajectories. Based on the spectroscopic observations, such a limitation certainly occurs. With a limitation on ionization near I_0 , any neutral species with a velocity distribution that peaks below the escape velocity and falls off rapidly enough with increasing velocity will form the inward peak.

Some workers^{13,16} believe that the thermal spike distribution we have assumed is accurate only for high energy (keV-MeV) sputtering ions. If neutrals are sputtered by corotating ions, a collision cascade velocity distribution of the form

$$f(v) dv \propto v^3 dv / (v^2 + v_0^2)^3 \quad (8)$$

may give a better representation. Here v_0 is the velocity corresponding to the binding energy. A relatively flat velocity distribution for neutrals does not form an inward peak if sources are distributed uniformly around I_0 's equator. The steepest velocity profile allowed by equation (8) corresponds to $v_0 = 0$; in Fig. 4b we have plotted the ion distribution obtained for this limiting case. As can be seen in Fig. 4b, a small peak at I_0 is formed. Even in the limiting case, a collision cascade distribution does not fall off rapidly enough to produce the required inward peak.

The results shown in Fig. 4 were calculated for sources independent of I_0 longitude, but there are reasons why ejection of particles from I_0 may be asymmetrical in longitude. For example, the nature of the plasma flow around I_0 may lead to preferential regions of access for the sputtering ions²⁴. This asymmetry would be highly dependent on the character of the interaction between I_0 and the torus plasma, which at this time is not well understood. An asymmetry in the source locations of sputtered particles could also occur if the dominant source of the neutrals on the surface were localized. SO_2 is found at roughly all longitudes; however, sulphur is confined to the region from 0 to 120° as illustrated in Fig. 2a (ref. 25). As mentioned previously, recent experiments²¹ have shown that sulphur gives very high sputtering yields, orders of magnitude greater than the sputtering yields estimated for SO_2 . Such large sputtering yields suggest that sulphur may be the dominant neutral injected into the torus. Furthermore, some calculations²⁶ suggest that a 4:1 injection ratio of sulphur to oxygen is consistent with the observed ion partitioning. If we consider sulphur neutrals with the thermal spike velocity distribution coming from the sulphur-rich regions on I_0 , a very sharp peak is formed (see Fig. 5a). Even for a collision cascade velocity distribution, the required peak is still formed if the sources are localized in the sulphur-rich regions (Fig. 5b).

Discussion

Our results show that the requirement that neutrals sputtered from I_0 must ultimately form a distribution of ions with a peak in NL^2 near $5.7 R_J$, substantially constrains models of the source. Ionization of neutrals on trapped trajectories must not contribute significantly to the ion population. The velocity distribution of ejected neutrals must fall off sufficiently rapidly with increasing velocity so as to favour heavily those particles near the escape velocity, or the source must be localized in longitude, or both. These requirements are easily reconciled with sulphur being the dominant neutral injected into the torus. S_2 sputtered from the surface of I_0 must first be dissociated before it is ionized, limiting the ionization of atomic S near I_0 . S_2^+ ions that are formed will be lost rapidly by dissociative recombination. The sulphur-rich volcanoes are, in fact, confined to longitudes favourable for producing the observed peak²⁵.

The high sputtering yields for sulphur²¹ allow a velocity distribution that falls off rapidly with increasing velocity and yet still gives an adequate source rate for maintaining the torus. For example, for a 3% escape fraction, 20% surface coverage and incident H^+ ions with energies $>220 \text{ keV}$, it has been estimated previously¹⁰ that a sputtering yield of one would give 10^{22} neutrals per second injected into the torus. A 4% escape fraction is obtained for the thermal spike velocity distribution of equation (2) with $T = 4,000 \text{ K}$. The measured sputtering yield of 10^6 for sulphur would then imply an injection rate of $>10^{28} \text{ s}^{-1}$, approximately the rate required from other arguments^{10,18,20}. This rate estimate may be low, as bombardment by heavier ions could lead to even higher sputtering yields, and lower energy ions (corotation energy to 220 keV) will undoubtedly contribute significant additional sputtering. Because flux-tube interchange diffusion produces strong outward mixing, the proposed sulphur source would produce a hot outer torus dominated by sulphur ions. It is not known whether sputtering of sulphur at I_0 will be as efficient as in the laboratory. In particular, the sputtering efficiency seems to increase with increasing source temperature²¹. In this regard, the 'hotspots' seen on I_0 ²⁷ may be areas of more efficient sputtering. The

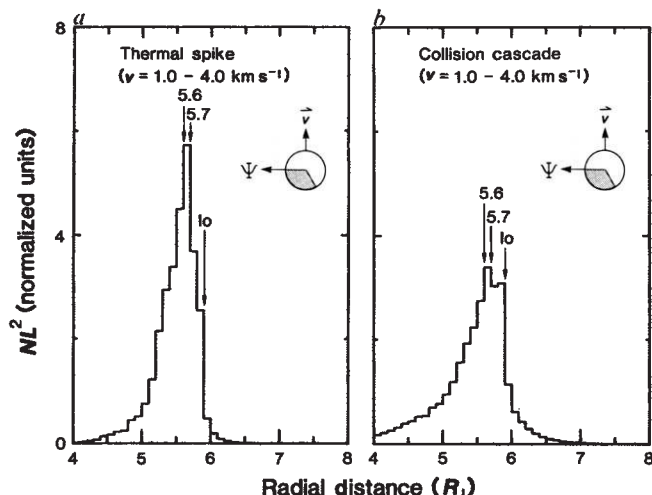


Fig. 5 *a*, As Fig. 4*a* except that the neutrals were ejected only from the sulphur-rich regions of Io²⁵ shown in Fig. 2*b*. A sharper inward peak is formed. As recent experiments²¹ indicate that sulphur gives anomalously high sputtering yields, the sulphur-rich regions may be the predominant source of Io torus ions. Ionization occurring in a narrow range of radial location near 5.7 R_J as shown here could explain the narrow, 'ribbon-like' structure seen in ground-based observations of sulphur ions¹. *b*, As *a*, except that a collision cascade velocity distribution with a binding energy of 0.1 eV was used. An adequate inward peak is still formed, even though the fall off of the velocity distribution is much more gradual. Recent experiments²¹ on the sputtering of sulphur have shown that for sputtering by 3-9 keV Ar⁺ ions, an effective surface binding energy of 0.1-0.3 eV is appropriate for sulphur.

longitudinal locations of the hotspots are also favourable for formation of the inward peak.

It is more difficult to reconcile our results with SO₂ as the dominant source. The requirement that the velocity distribution fall off rapidly with velocity, coupled with the much lower sputtering yields for SO₂, make it unlikely that a dominant SO₂ source on the surface of Io could provide the necessary escape rates to populate the torus. Although our results do not eliminate sputtering from a thick SO₂ (or sulphur) atmosphere as the dominant source, they do place constraints on this possibility. It has been noted^{10,28} that sputtering from such an atmosphere can easily maintain the required total source rate of neutrals. However, the velocity distribution expected for this source²⁸ (similar to a collision cascade distribution) does not fall off rapidly enough for the neutrals to form the inward peak unless the source is longitudinally asymmetrical. The primary sputtering product from a thick SO₂ atmosphere (density of $\sim 10^{10}$ cm⁻³ at the surface) would be atomic oxygen^{10,20}, so again a mechanism that limits the ionization of neutrals near Io would be necessary. If the source is a thick SO₂ atmosphere, sputtering models predict more than adequate oxygen escape rates but they encounter problems in obtaining even the minimum required sulphur escape rates^{10,20}. A thick atmosphere on the dayside of Io coupled with a thin atmosphere on the nightside that allows sputtering of sulphur from the surface of Io could resolve this dilemma.

In our calculations, assumptions regarding lifetimes for neutrals ignore any variation of ionization probability with position. Thus, we have effectively assumed a uniform electron density (n_e) and temperature (T_e) distribution throughout the torus. In actuality, n_e and T_e drop sharply inside 5.7 R_J , making ionization in the region of the observed peak more likely. By using a model with uniform n_e and T_e , we are able to investigate the question of whether or not the peak can form inward of Io solely because neutral particles tend to spend most of their time in this region. As our calculations have shown, there are physically plausible scenarios in which neutral particles will tend to be ionized to form a peak inwards of Io, even for uniform

distributions of n_e and T_e . This then explains how the peak in the ion density could form inwards of Io in the first place. Heating of electrons by collisions with pickup ions would then, in turn, produce the observed radial variation of T_e (ref. 29). The T_e variation would increase the probability of ionization near the peaks of the derived profiles of newly ionized neutrals. Thus, in a self-consistent calculation, the locations of the peaks we have obtained would not change, but their steepness would be enhanced. We believe that 'the ribbon' of S(II) emission observed¹ near 5.7 R_J may be explained in this way. We have not taken into account the possible effects of collisions on the trajectories in our calculations. Although collisional processes could introduce some spread into the trajectories of the neutrals, our basic result, that the inward peak forms if the distribution of neutral velocities falls off rapidly with increasing velocity, should remain unaffected.

By assuming that the dominant neutral source of torus ions is S₂, we showed that only small numbers of ions are created near Io as necessary for consistency with observations. The limitation to the ionization rate near Io can be achieved in other ways. For example, the presence of the neutral cloud near Io can limit the amount of ionization. Within 3 Io radii of Io's surface, the density of sputtered neutrals will be much greater than the electron density, so it is possible that ionization near Io is limited by the availability of the electrons or by electron cooling resulting from interaction with the neutral cloud. A cooler electron temperature in the Io flux tube has in fact been observed³⁰. If the presence of the neutral cloud does in fact limit ionization and other neutral species meet the constraints on the velocity distribution and/or source locations for formation of the inward peak, then S₂ need not be the dominant neutral injected into the torus.

Similar trajectory calculations have been performed previously for sodium^{31,32} and a model, including the electron density and temperature profiles, has been published for oxygen³³. The latter calculation provides useful insight into the supply of ions to the torus, but it does not explain why the radial profiles of ion density and NL^2 peak inwards of Io's orbit. In the cited calculation³³, the measured n_e and T_e profiles were assumed and then the position where the particles were ionized was determined. The calculated radial profile of newly created ions obtained in ref. 33 peaks roughly at Io's orbit (Figure 8 of ref. 33), farther out than the peak in the ion-density data. As electron and ion number densities must peak in the same location, the electron distribution assumed for the calculation in ref. 33 would not be maintained in a steady state. Furthermore, because the newly created pickup ions are responsible for the radial variations of n_e and T_e , a calculation that assumes observed radial profiles for these quantities does not demonstrate how the peaks developed inwards of Io in the first place. In addition, we note that the calculated³³ peak in the ion-creation rate was obtained for particles launched from the exobase (at 2,600 km) with a velocity of 2.6 km s⁻¹. This velocity maps roughly to a velocity of 3.0 km s⁻¹ for a particle leaving the surface of Io. As seen in Fig. 4*b*, when velocities much larger than the escape velocity become important, the inward peak does not form for sources distributed uniformly in longitude. In fact, for a velocity of 3.0 km s⁻¹, even a source restricted in longitude will not form the inward peak. Thus it seems that the neutral velocity chosen for the calculations of ref. 33 is too high to form a peak much inside Io's orbit. Finally, we note that because neutrals with velocities below the escape velocity were not considered in ref. 33, the important effects of ionization of neutrals on trapped trajectories were ignored in their calculations.

Conclusions

We have investigated models for the formation of the peak in ion density and NL^2 inwards of Io's orbit. We summarize our findings as follows: (1) Neutral particles with velocities above the escape velocity sputtered from sources independent of Io longitude will form a peak inward of Io between 5.6 and 5.7

R_1 , provided that the velocity distribution drops off sufficiently rapidly. We have shown that a thermal spike velocity distribution (equation (2)) produces such a peak whereas distributions that fall off as v^{-3} or less do not create the required peak. (2) If the sputtering sources are confined to longitudinal regions on Io where sulphur predominates²⁵, the inward peak is obtained even for less steeply decreasing velocity distributions (such as collision cascade). If a thermal spike velocity distribution is used, a sharper inward peak is formed. (3) Velocity distributions predicted for sputtering of neutrals place many more particles on trapped (non-escape) trajectories than on escape trajectories, so an adequate source rate²⁰ of ions in the torus ($2-7 \times 10^{28} \text{ s}^{-1}$) implies that a dense neutral cloud ($10^5-10^6 \text{ cm}^{-3}$) will form within 3 Io radii of Io's surface. Ionization of atomic sulphur and oxygen in this cloud must be limited by some process to satisfy observational constraints on ionization¹⁷ and for formation of the inward peak. (4) The above results are more readily reconciled with S_2 being the dominant neutral injected into the torus rather than SO_2 or O. In particular, the large sputtering yields recently demonstrated for sulphur²¹ would allow sputtered sulphur to have a steeply decreasing velocity distribution and still give an adequate source rate of neutrals into the torus. In addition, the sulphur-rich volcanoes on Io are confined to a longitudinal region favourable for producing the inward peak²⁵.

A dominant SO_2 source on the surface is unlikely, because if SO_2 were sputtered with the required rapidly decreasing velocity distribution, the source rate of neutrals into the torus would seem to be inadequate. If the presence of the dense neutral cloud near Io limits ionization of all neutral species, a thick SO_2 atmosphere cannot be eliminated; however, velocity distributions previously estimated for this possibility²⁸ do not fall off rapidly enough with increasing velocity to produce the inward peak without invoking a source that is asymmetrical in longitude. (5) Although the above arguments suggest that the dominant source of ions in the hot torus is neutral sulphur, consistent with previous evidence from ion partitioning²⁶, an SO_2 or O source with a collision cascade velocity distribution would provide ions more widely distributed in radial distance and these ions may dominate the inner cool torus, whose composition need not be the same as that of the hot torus.

We thank F. Bagenal and J. T. Trauger for providing advance copies of their recent papers on the Io torus. We are grateful to P. R. Higbie for calling our attention to recent work on sputtering of sulphur and to W. H. Smyth for encouraging us to reconsider assumptions regarding the distribution of lifetimes. This work was partially supported by NASA under grant NSG-7295 and by the Jet Propulsion Laboratory under contract 955232.

Received 28 January; accepted 11 March 1985.

1. Trauger, J. T. *Science* **226**, 337-341 (1984).
2. Bagenal, F. J. *geophys. Res.* **90**, 311-324 (1985).
3. Bagenal, F. & Sullivan, J. D. *J. geophys. Res.* **86**, 8447-8466 (1981).
4. Richardson, J. D., Siscoe, G. L., Bagenal, F. & Sullivan, J. D. *Geophys. Res. Lett.* **7**, 37-40 (1980).
5. Siscoe, G. L. & Summers, D. J. *geophys. Res.* **86**, 8471-8479 (1981).
6. Siscoe, G. L. *et al.* *J. geophys. Res.* **86**, 8480-8484 (1981).
7. Richardson, J. D. & Siscoe, G. L. *J. geophys. Res.* **86**, 8487-8490 (1981).
8. Brown, R. A. & Ip, W. H. *Science* **213**, 1493-1495 (1981).
9. Durrance, S. T., Feldman, P. D. & Weaver, H. A. *Astrophys. J.* **267**, 1125-1129 (1983).
10. Cheng, A. F. J. *geophys. Res.* **89**, 3939-3944 (1984).
11. Hill, T. W., Dessler, A. J. & Michel, F. C. *Geophys. Res. Lett.* **1**, 3-6 (1974).
12. Peach, G. J. *Phys.* **B4**, 1670-1677 (1971).
13. Brown, R. A., Shemansky, D. E. & Johnson, R. E. *Astrophys. J.* **264**, 309-323 (1983).
14. Brown, R. A., Pilcher, C. B. & Strobel, D. F. in *Physics of the Jovian Magnetosphere* (ed. Dessler, A. J.) 197-225 (Cambridge, New York, 1983).
15. Watson, C. C. *Proc. Lunar planet. Sci.* **12b**, 1569-1583 (1981).

16. Tombrello, T. A. *Radiation effects* **65**, 149-158 (1982).
17. Shemansky, D. E. *Astrophys. J.* **242**, 1266-1277 (1980).
18. Hill, T. W. *Science* **207**, 301-302 (1980).
19. Pontius, D. H. & Hill, T. W. *Geophys. Res. Lett.* **9**, 1321-1324 (1982).
20. Kumar, S. J. *geophys. Res.* **89**, 7399-7406 (1984).
21. Fink, D. *et al.* *Nucl. Instr. Meth. Phys. Res.* **B1**, 275-281 (1984).
22. Cheng, A. F. J. *geophys. Res.* **87**, 5301-5304 (1982).
23. Walls, F. L. & Dunn, G. H. J. *geophys. Res.* **79**, 1911-1915 (1974).
24. Hill, T. W., Dessler, A. J. & Fanale, F. P. *Planet. Space Sci.* **27**, 419-424 (1979).
25. McEwen, A. S. & Soderblom, L. A. *Icarus* **55**, 191-217 (1983).
26. Moreno, M. A., Newman, W. I. & Kivelson, M. G. *J. geophys. Res.* (submitted).
27. Johnson, T. V. *et al.* *Science* **226**, 134-237 (1984).
28. Haff, P. K., Watson, C. C. & Yung, Y. L. *J. geophys. Res.* **86**, 6933-6938 (1981).
29. Barbosa, D. D., Coroniti, F. V. & Eviatar, A. *Astrophys. J.* **274**, 429-442 (1983).
30. Sittler, E. C. & Strobel, D. F. *EOS Trans.* **65**, 1038 (1984).
31. Smyth, W. H. & McElroy, M. B. *Planet. Space Sci.* **25**, 415-431 (1977).
32. Smyth, W. H. *Astrophys. J.* **204**, 708-725 (1983).
33. Smyth, W. H. & Shemansky, D. E. *Astrophys. J.* **273**, 865-875 (1983).

Speed of sound in the solar interior

J. Christensen-Dalsgaard*, T. L. Duvall Jr[†], D. O. Gough[‡], J. W. Harvey[§] & E. J. Rhodes Jr^{||}

* NORDITA, Blegdamsvej 17, København Ø, Denmark and Department of Applied Mathematics and Theoretical Physics, Silver Street, Cambridge CB3 9EW, UK

[†] Laboratory for Astronomy and Solar Physics, NASA/Goddard Space Flight Center, Greenbank, Maryland 20771, USA

[‡] Institute of Astronomy & Department of Applied Mathematics and Theoretical Physics, Madingley Road, Cambridge CB3 0HA, UK

[§] National Solar Observatory, National Optical Astronomy Observatories, Tucson, Arizona 85726, USA

^{||} Department of Astronomy & Space Sciences Center, University of Southern California, Los Angeles, California 90007, USA

Frequencies of solar 5-min oscillations can be used to determine directly the sound speed of the solar interior. The determination described here does not depend on a solar model, but relies only on a simple asymptotic description of the oscillations in terms of trapped acoustic waves.

SOLAR 5-min oscillations are acoustic waves that propagate in a waveguide beneath the surface of the Sun. For any particular wave, the guide is a spherical shell whose boundaries depend on the frequency and wavenumber of the wave. Upward directed waves are reflected just beneath the photosphere, where the density scale height is comparable with the wavelength. Refraction of a wave initially propagating downwards deflects the path back towards the surface, at a point where the local horizontal component of the speed of the wave is equal to the sound speed.

The frequency of a wave that resonates in the waveguide depends on the variation of the sound speed with depth. The latter determines both the locations of the boundaries of the

cavity within which the wave can propagate and the time it takes for the wave to traverse that cavity. By considering successively the frequencies of a sequence of modes that penetrate deeper and deeper, it is possible, at least in principle, to measure how the sound speed varies with depth.

Using an approximation to the observed dispersion relation presented in ref. 1 for modes with degrees in excess of 100, it has already been possible to estimate the sound speed in the upper half of the convection zone². More recent observations^{3,4} have isolated modes with degrees as low as unity, which sample almost the entire volume of the Sun. Therefore, we should now have sufficient data to infer the sound speed almost everywhere.

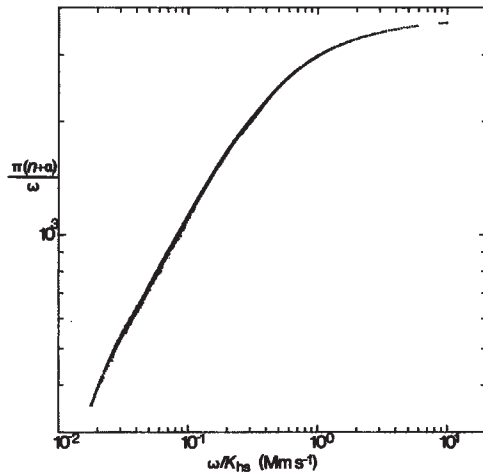


Fig. 1 Observed modes of solar 5-min oscillations collapsed according to the dispersion relation (1) with $\alpha = 1.58$. Units of ω and k_{hs} are s^{-1} and $Mm s^{-1}$, respectively.

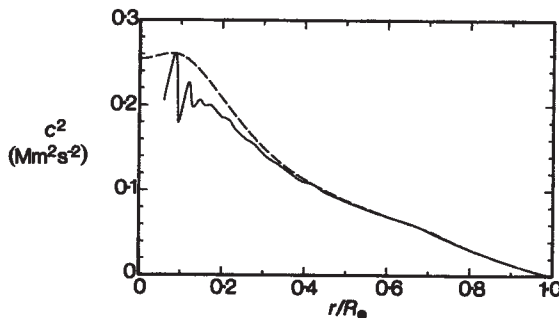


Fig. 2 Result of inverting computed eigenfrequencies of a solar model, using equation (13) to obtain an estimate c_m^2 (continuous line) of the square of the sound speed. Dashed line, square of the actual sound speed c^2 in the model. Units of speed, $Mm s^{-1}$.

Dispersion relation

A convenient form for analysing high-frequency p modes was introduced by Duvall¹. He showed that if $\pi(n + \alpha)/\omega$ is plotted against ω/k_{hs} , where n is the order of the mode, ω is its frequency and k_{hs} is the horizontal component of its wave number vector in the photosphere, a value of the constant α could be found for which high-degree 5-min modes lie very close to a single curve. Thus,

$$\frac{\pi(n + \alpha)}{\omega} \approx F(\omega/L) \quad (1)$$

where L is related to the degree l of the mode by

$$L^2 = l(l+1) = R_\odot^2 k_{hs}^2 \quad (2)$$

$R_\odot$ being the radius of the Sun. For the modes considered originally, $\alpha = 1.50$. This relation, with a somewhat different value of α , holds also for the eigenfrequencies of a solar model². It is exact for a plane-parallel adiabatically stratified polytropic envelope^{6,7} and deviations from it estimate an upper bound to the departure of the temperature gradient in the solar convection zone from its adiabatic value². We now demonstrate that the relation is an asymptotic property of high-order acoustic modes in a spherical star, even when the stratification is not polytropic.

With respect to spherical polar co-ordinates (r, θ, ϕ) the vertical component v of velocity, say, of a linearized normal mode, has the functional form

$$v(r, \theta, \phi, t) = \text{Re}[V(r)Y_l^m(\theta, \phi)e^{-i\omega t}], \quad (3)$$

where Y_l^m is a spherical harmonic of degree l and order m . Similar expressions hold for the perturbations in pressure and

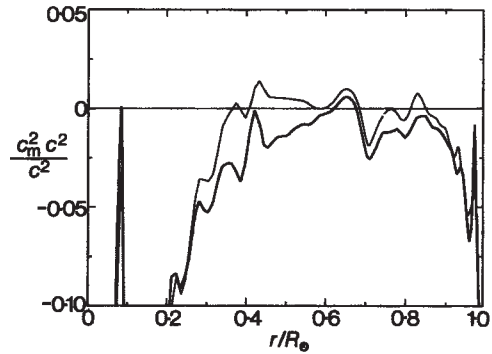


Fig. 3 Thick line, relative deviation $(c_m^2 - c^2)/c^2$ of the inferred squared sound speed c_m^2 shown in Fig. 2 from the actual value of c^2 in the theoretical model. Thin line, relative deviation from c^2 of the squared sound speed c_0^2 in Fig. 4 inferred from the solar data.

density. The amplitude function V satisfies a differential equation which admits, with appropriate boundary conditions, a discrete set of eigenfunctions that can be labelled with an integer n . The problem can be treated by standard asymptotic techniques when n is large⁸⁻¹⁰. The following simplified argument contains the essential physics, and to a good approximation yields the correct dispersion relation.

For modes with frequencies well in excess of the maximum buoyancy frequency in the Sun's interior (~ 0.5 mHz in standard theoretical solar models) the effect of buoyancy may be neglected. Furthermore, high-order modes produce a negligible perturbation to the gravitational potential. They may be characterized by a local wavenumber k whose vertical component $k_v(r)$ is much greater than the curvature and the inverse scale heights of the unperturbed structure of the Sun. The waves are therefore locally plane, satisfying

$$v(r, \theta, \phi, t) \propto e^{i(kr - \omega t)} \quad (4)$$

where r is the position vector and k and ω are related by the pure acoustic dispersion relation

$$\omega^2 = k^2 c^2 \equiv (k_v^2 + k_h^2) c^2 \quad (5)$$

with $k = |k|$. This simple relation is not valid close to the upper reflecting boundary of the cavity, where scale heights are comparable with k_v^{-1} . Nor indeed does it even predict the existence of that reflecting boundary. Where the relation is valid, however, it determines $k_v(r)$ according to

$$k_v^2 = k^2 - k_h^2 = \frac{\omega^2}{c^2} - \frac{L^2}{r^2} \quad (6)$$

Near the solar surface $\omega^2/c^2 \gg L^2/r^2$, unless L is very large, and $k_v^2 \gg k_h^2$. The waves propagate almost vertically, and the structure of the amplitude function V is essentially independent of the horizontal component k_h of the wave number vector. As r decreases, so does ω^2/c^2 . Thus, according to equation (5), k_v^2 decreases with depth. The wave is refracted away from the vertical, and is eventually turned back at $r = r_t$, where

$$r_t^2/c^2(r_t) = L^2/\omega^2 \quad (7)$$

and propagation is locally horizontal.

From a comparison of the separated form of equation (3) with the local approximation in equation (4), it is evident that beneath the outer reflection point $V(r)$ is an oscillatory function where $r > r_t$. In particular, the difference in phase between the inner turning point $r = r_t$ and the outer reflection point $r = R_t$ is

$$\Delta\Psi = \int_{r_t}^{R_t} k_v dr \quad (8)$$

Had the mode been trapped between rigid surfaces the condition imposed by the waveguide would have been $\Delta\Psi = n\pi$, where n

is an integer. But the existence of evanescent regions above and below the cavity introduces an additional phase shift $\pi\alpha'$. Therefore, the dispersion relation is

$$\Delta\Psi = \int_{r_i}^{R_\odot} \left(\frac{\omega^2}{c^2} - \frac{L^2}{r^2} \right)^{1/2} dr = \pi(n + \alpha') \quad (9)$$

Because the upper limit of integration is close to the surface of the Sun, we replace it by $R_\odot$ and absorb the small change in the integral into α' , dropping the prime. The integral is now the product of ω and a function of ω/L alone, because, according to equation (7), r_i depends on ω and L only in the combination ω/L . Consequently the dispersion relation (9) becomes

$$F(w) \equiv \int_{r_i}^{R_\odot} \left(\frac{r^2}{c^2} - \frac{1}{w^2} \right)^{1/2} \frac{dr}{r} = \frac{\pi(n + \alpha)}{w} \quad (10)$$

where $w = \omega/L$.

The phase shift introduced at the lower turning point is only weakly dependent on the structure of the solar model, because there the local dispersion relation (5) is a good approximation and the geometry of the deflection of the wave is adequately taken into account by the integrand in equation (10). Indeed, equation (9) may be derived directly from ray theory¹¹, in which $\Delta\Psi$ is expressed as an integral along a ray whose geometry is determined by equation (6).

For low-degree modes $r_i/R_\odot \ll 1$ and equation (10) may be approximated by

$$\omega \approx \pi(n + \frac{1}{2}L + \alpha) \left(\int_0^{R_\odot} \frac{dr}{c} \right)^{-1} \quad (11)$$

This can be compared with the dispersion relation for high-order acoustic modes of an isothermal gas confined in a rigid sphere, which is known exactly, and whose leading term is of the form in equation (11) with $L + 2\alpha$ replaced by $l + 1$. Thus, in this limit the form of equation (10) seems to be roughly correct, provided l is not too small.

To estimate the phase factor α it is necessary to account also for the phase shift on reflection at $r = R_i$. For this we consider the limit of high degree. In that case the modes are confined to the convection zone, which we approximate by an adiabatically stratified plane-parallel polytrope of index μ under constant gravity g . Then equation (10) becomes

$$\omega^2 = 2\mu^{-1}(n + \alpha)gk_{\text{hs}}, \quad (12)$$

which is identical to the exact dispersion relation under these conditions^{6,7} if $\alpha = \mu/2$. Since conditions near $r = R_i$ are essentially independent of l , we deduce that this constant value of α is approximately valid for all l . Then equation (10) takes the form of equation (1). Moreover, if L were replaced by $l + \frac{1}{2}$, the approximation (11) would agree precisely with the leading term of previous asymptotic analyses^{8,10} of low-degree p modes.

Observations

Measurements of $\omega(n, l)$ were made from two sets of Doppler observations using one or more of three different methods. The first data set¹ consists of a spectrum of prograde and retrograde sectoral and nearly sectoral modes with $105 \leq l \leq 892$. The resolution is ~ 10 in l and the uncertainty in the spatial scale is $\sim 0.5\%$. Frequencies were estimated by centroid analysis, and zonal-mode frequencies were deduced by averaging corresponding prograde and retrograde frequencies. The second data set⁴ consists of a spectrum of zonal oscillations with $1 \leq l \leq 200$, with a resolution of ~ 3 in l and a scale uncertainty of $\sim 0.5\%$. For $l \leq 20$ individual modes could be isolated and identified, and their frequencies measured to $\pm 2 \mu\text{Hz}$. For $l > 20$ spline functions were fitted by eye to visual representations of sets of modes of like order n and used to determine the frequencies; we estimated that the frequencies determined in this way are correct

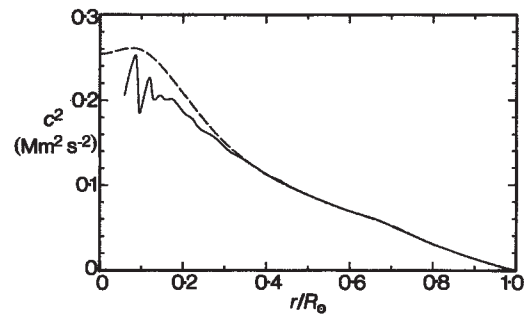


Fig. 4 Square of the interior sound speed c_0^2 in the Sun (continuous line) obtained by inverting the data in Fig. 1. The value of c^2 in solar model is included as a dashed line for comparison.

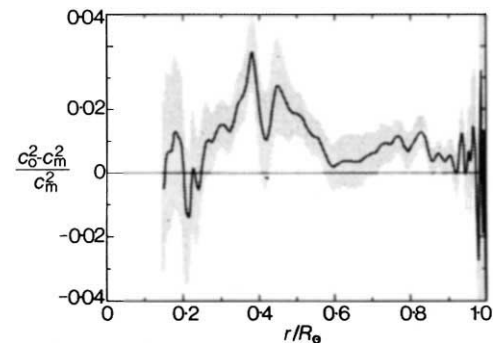


Fig. 5 Relative discrepancy $(c_0^2 - c_m^2)/c_m^2$ between the squared sound speeds in Figs 2 and 4 obtained by inverting corresponding oscillation frequencies of the Sun and the solar model. The shaded region lies within one standard deviation of that part of the error in the inferred values of c_0^2 resulting from random errors in the data. It was computed from many theoretical realizations of the data obtained by adding random errors to the real solar data with the same variance as the error estimates that are indicated in the text.

to within $\pm 10 \mu\text{Hz}$. When a mode was present in both data sets the average frequency was adopted in the subsequent analysis. A comparison of the original data sets revealed a systematic discrepancy that was larger than the estimated errors in either. The root of the discrepancy was traced to an error in the reduction of the high-degree data, which caused spatial scales to have been overestimated by 1%. Accordingly we increased the values of l in the original high-degree data set by 1%, although an increase of 1.4% would have minimized the discrepancy. In total, 2,820 modes were used, with lower turning points r_i in the range $0.06\text{--}0.994 R_\odot$. Note that the quality of the estimates and the degree of smoothing changes at $l = 20$ and in the range $105\text{--}200$.

The relation defined by equation (1) was fitted to the combined data, which are shown in Fig. 1. It was determined by choosing α such as to minimize the spread in the data, spread being measured by the sum of the squares of the altitudes of all triangles whose vertices are triplets of data points in Fig. 1 with consecutive values of ω/k_{hs} ; for this purpose the base of a triangle is defined as the line joining the two points of a triplet with the extreme values of ω/k_{hs} . This yields $\alpha = 1.58$; other definitions of spread, such as the r.m.s. deviations of the data from an appropriately chosen smooth curve, gave similar results.

On the whole the scatter in the data decreases as ω/k_{hs} increases: those modes considered in ref. 1 having $\omega/k_{\text{hs}} \leq 0.1 \text{ Mm s}^{-1}$ have an r.m.s. deviation from a smooth curve of $\sim 0.4\%$, whereas the deviation is less than 0.1% when $\omega/k_{\text{hs}} > 1 \text{ Mm s}^{-1}$. Theoretical eigenfrequencies of a solar model behave similarly. Therefore, it seems that much of the variation can be

attributed to the inaccuracy of our simple asymptotic analysis, which increases as l increases.

Inversion of dispersion relation

The first part of equation (10) can be inverted¹¹ to yield r as a function of c/r :

$$r = R_{\odot} \exp \left\{ -\frac{2}{\pi} \int_{c/r}^{c/r} [w^{-2} - (c/r)^{-2}]^{-1/2} \frac{dF}{dw} dw \right\} \quad (13)$$

where c_* is the value of c at $r = R_{\odot}$. According to the second part of equation (10), the function F is equal to the observed quantity $\pi(n + \alpha)/\omega$. Consequently equation (13) can be used to determine $c(r)$ from the observations presented in Fig. 1 without any recourse to a solar model.

Given the simplified nature of the arguments leading to equation (10), the reliability of equation (13) may at first sight seem rather dubious. Therefore before using it to analyse the real solar data we tested it on a theoretical solar model. Adiabatic eigenfrequencies of the modes corresponding to those that had been observed on the Sun were computed, using the equilibrium solar Model 1 of ref. 12. These were fitted to the relation (1) in the manner described above, yielding $\alpha = 1.39$. A representation of the function F was then obtained by applying twice a 15-point running mean to the theoretical analogue of Fig. 1. The smoothing is necessary because deviations from the simple asymptotic law of equation (10) cause modes of different order to lie on somewhat different curves. The resulting estimate of $F(w)$ was extrapolated to $w = c_*/R_{\odot}$ with the power law Aw^a , with the coefficients $a = 1.035$ and $A = 1.70 \times 10^7$ (in c.g.s. units) chosen to match smoothly onto the lower end of the theoretical data. Its derivative was evaluated by centred second-order accuracy differencing of the smoothed representation of the data. Finally the expression on the right-hand side of equation (13) was evaluated using a difference scheme tailored to accommodate the square-root singularity in the integrand.

In Fig. 2 the square of the sound speed (c_m^2) inferred by this procedure is compared with the actual value of c^2 in the solar model. In a substantial region the result of the inversion is quite accurate. This is emphasized in Fig. 3, which shows the relative errors in the inferred values of c^2 . For $r/R_{\odot}$ between 0.4 and 0.9, the error in c^2 is $<2.5\%$.

In view of the breakdown of the assumptions of the asymptotic theory near the surface, the large error when $r > 0.9 R_{\odot}$ is not surprising. The reason for the errors in the deep interior is less obvious, although it appears to arise in part from the fact that the inversion procedure is incremental, and can accumulate errors in the asymptotic analysis that occur nearer the surface.

Determination of solar sound speed

The inversion of the solar data in Fig. 1 was performed in exactly the same way as for the artificial data, except that the parameters defining the power law used to extrapolate F to the surface were $a = 0.846$, $A = 2.63 \times 10^6$. The results are presented in Figs 3–5.

The qualitative features of the inversions shown in Figs 2 and 4 are quite similar. A quantitative comparison is made in Fig. 5, in which the relative deviation of the square of the inferred solar sound speed (c_0^2) from that inferred by inverting the eigenfrequencies of the theoretical model (c_m^2) is plotted. In the region between $0.4 R_{\odot}$ and $0.9 R_{\odot}$, where the inversion is expected to work well, c_0^2 is within 3% of c_m^2 . Furthermore, it is striking how the difference between c_0^2 and c_m^2 follows the errors in the inversion of the theoretical data, both of which are shown in Fig. 3. In particular, the behaviour is similar above $0.9 R_{\odot}$ and below $0.4 R_{\odot}$, where our inversion technique seems to fail. Systematic errors arising from inaccuracies in the simple asymptotic procedure are similar for the Sun and the solar model, suggesting that the error in Fig. 5 is substantially less than the discrepancies plotted in Fig. 3.

To test this hypothesis, we compared inversions of 5-min eigenfrequencies of two theoretical models of the Sun: Model

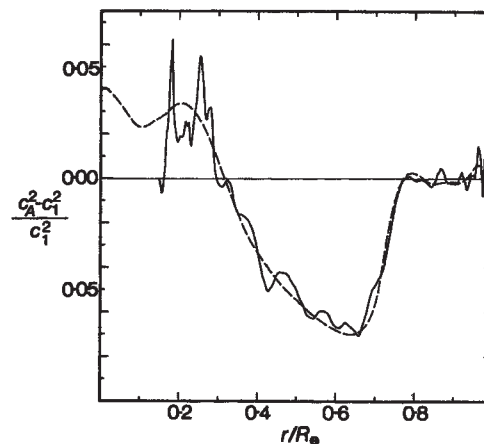


Fig. 6 Solid line is the relative difference $(c_A^2 - c_1^2)/c_1^2$ between the inferred squared sound speeds c_1^2 and c_A^2 of Model 1 and Model A of the Sun obtained by inverting the frequencies of 363 5-min modes. Dashed line, actual relative difference.

A of Christensen-Dalsgaard *et al.*¹³ and Model 1 (see ref. 12). The sound speeds in the two models differ substantially beneath the convection zone and thus provide a basis for assessing the significance of the deviations plotted in Fig. 5. Inversions were performed on 363 common modes whose frequencies lie in the range spanned by the solar observations. The relative difference between the squares of the sound speeds thus inferred is plotted in Fig. 6, together with the actual difference. Because relatively few modes were used, the errors in the inversions are rather greater than those illustrated in Fig. 3. However, they largely cancel in the difference plotted in Fig. 6, suggesting that except close to the solar surface and close to the centre, inversion errors contribute less to Fig. 5 than errors in the data. We conclude also from this comparison that the inversion technique is not sensitive to the detailed structure of the surface layers of the Sun. The sound speeds in Model A and Model 1 differ by up to 8% near the upper turning point R_t , and influence the frequencies of all 5-min modes whatever the value of l . However, as Fig. 6 demonstrates, those influences are successfully cancelled in that aspect of the function $F(w)$ that determines the sound speed at greater depths. This is achieved largely by absorbing the influence of the surface layers into the constant α , which takes different values for the two solar models.

We have also determined the effect of the spatial scale error in the original high-degree data by comparing inversions with and without the 1.01 scale factor imposed on l . At radii greater than $0.8 R_{\odot}$ the inferred sound speeds differed by $\sim 0.5\%$. At greater depths the difference decreased, and was $<0.1\%$ for $r < 0.6 R_{\odot}$.

Discussion

It seems that in the region where we have succeeded in measuring $c(r)$ the differences between the solar values and those of the theoretical model are quite small. However, there is an indication that for r between $0.3 R_{\odot}$ and $0.5 R_{\odot}$ the sound speed in the Sun is generally somewhat greater than it is in the model. Deviations of the inferred sound speeds from a smooth curve are clearly similar in the two cases, and are partly dependent on mode selection. An earlier direct comparison between observed and calculated oscillation frequencies¹⁴ suggested that there is an error in Model 1 somewhere between the turning points of the 5-min modes with $l = 20$ and with $l = 40$, which we find are located at $r/R_{\odot} \approx 0.45$ and 0.70 respectively.

A change of curvature can clearly be seen in Figs 2 and 4 at $r \approx 0.7 R_{\odot}$. This indicates the position of the base of the convection zone, which in the theoretical solar model is at $r = 0.71 R_{\odot}$. In this respect the theory appears to mimic reality, and we deduce that the convection zone in the Sun is ~ 200 Mm deep. This value agrees with previous model-fitting calibrations by Berthomieu *et al.*¹⁵ and by Shibahashi *et al.*¹⁶, and is somewhat

greater than the value for the solar model preferred by Ulrich and Rhodes^{17,18}.

The relatively high sound speed deduced near $r=0.4 R_{\odot}$ deserves comment. A possible explanation is that the opacity used to construct the theoretical solar model is too low in the outer part of the radiative interior. An increase of $\sim 20\%$ in opacity at temperatures between $\sim 10^5$ and $\sim 4 \times 10^6$ K, for example, would reduce the discrepancy to an insignificant level.

It has not yet been possible to determine the sound speed in the energy-generating core. This is particularly disappointing, for we had hoped to ascertain whether there had been substantial mixing in the chemically inhomogeneous layers. There are certainly grounds for believing, however, that the information is contained in the data¹⁹.

It is of interest that the values of α obtained from the solar model and from the observations, namely 1.39 and 1.58, are significantly different, which indicates that errors in the theory of the surface layers have been made, either in the construction of the equilibrium model or in the oscillation calculation. This has also been found from comparing the observed and calculated frequencies directly¹⁴. A strength of the present approach is that it allows one to separate, in α , the effects of the surface layers where the physics of the oscillations is incompletely understood. The frequencies computed by Berthomieu *et al.*¹⁵ and Shibahashi *et al.*¹⁶ have $\alpha = 1.53$ and 1.51 respectively; the corresponding values of α computed from just those observed modes that are common with the calculations are 1.55 and 1.54. Thus, these

models seem to represent the surface of the Sun more accurately. The frequencies published by Ulrich and Rhodes¹⁸ have $\alpha = 1.40$, the corresponding value from the observations being 1.58.

Solar oscillation frequencies do not satisfy the relation in equation (1) exactly. The deviations, although small, are significant compared with the accuracy of the observations. Thus, we have not fully used the information contained in the data. To do so one must use either a higher-order asymptotic approach or a direct numerical calculation of the frequencies. In either case inversion would almost certainly be an iterative process. In general the result of such an analysis, like many a solution of a nonlinear problem, might depend on an initial guess of a model and may not necessarily be unique. On the other hand the JWKB asymptotic analysis used here does provide, in equation (13), a unique determination of the sound speed from the observed function F . The asymptotic analysis is valid if the scale of variation of the background state is much greater than all oscillation wavelengths considered, and that condition is indeed satisfied throughout most of the interior of any traditional theoretical model of the Sun. Therefore it is plausible that by using this asymptotic determination as an initial trial in a more accurate inversion, and imposing constraints of smoothness on the solution resulting from the iteration, a good model representing the large-scale structure of the Sun that satisfies the observed frequencies might be determined.

We thank S. Frandsen for useful conversations and R. K. Ulrich for helping us to improve the presentation.

Received 21 November 1984; accepted 27 March 1985.

1. Duvall, T. L. Jr *Nature* **300**, 242-243 (1982).
2. Gough, D. O. *Mem. Soc. astr. Italiana* **55**, 13-35; in *Solar Physics and Interplanetary Travelling Phenomena* (eds Chen, B. & de Jager, C.) (Yunnan Observatory, Kunming, 1984).
3. Duvall, T. L. Jr & Harvey, J. W. *Nature* **302**, 24-27 (1983).
4. Harvey, J. W. & Duvall, T. L. Jr *Proc. Conf. solar Seismol. Space* (eds Ulrich, R. K. *et al.*) 165-172 (Jet Propulsion Laboratory Publication 84-84, Pasadena, 1984).
5. Christensen-Dalsgaard, J. & Gough, D. O. *Bull. Am. astr. Soc.* **13**, 859-860 (1981).
6. Gough, D. O. *Proc. IAU Colloq.* No. 36, 3-36, 1977.
7. Christensen-Dalsgaard, J. *Mon. Not. R. astr. Soc.* **190**, 765-791 (1980).
8. Vandakurov, Yu. V. *Astr. Zh.* **44**, 786-797 (1967).
9. Unno, W., Osaki, Y., Ando, H. & Shibahashi, H. *Nonradial Oscillations of Stars* (University of Tokyo Press, 1979).

10. Tassoul, M. *Astrophys. J. Suppl.* **43**, 469-490 (1980).
11. Gough, D. O. *Phil. Trans. R. Soc. A* **313**, 27-38 (1984).
12. Christensen-Dalsgaard, J. *Mon. Not. R. astr. Soc.* **199**, 735-761 (1982).
13. Christensen-Dalsgaard, J., Gough, D. O. & Morgan, J. G. *Astr. Astrophys.* **73**, 121-128; **79**, 260 (1979).
14. Christensen-Dalsgaard, J. & Gough, D. O. *Proc. Conf. solar Seismol. Space* (eds Ulrich, R. K. *et al.*) 199-204 (Jet Propulsion Laboratory Publication 84-84, Pasadena, 1984).
15. Berthomieu, G. *et al.* in *Nonradial and nonlinear stellar pulsation* (eds Hill, H. A. & Dziembowski, W. A.) 307-312 (Springer, Heidelberg, 1980).
16. Shibahashi, H., Noels, A. & Gabriel, M. *Astr. Astrophys.* **123**, 283-288 (1983).
17. Ulrich, R. K. *Astrophys. J.* **258**, 404-413 (1982).
18. Ulrich, R. K. & Rhodes, E. J. Jr *Astrophys. J.* **265**, 551-563 (1983).
19. Gough, D. O. *Phys. Bull.* **34**, 502-506 (1983).

Direct mutagenesis of Ha-ras-1 oncogenes by N-nitroso-N-methylurea during initiation of mammary carcinogenesis in rats

Helmut Zarbl, Saraswati Sukumar, Anne V. Arthur, Dionisio Martin-Zanca & Mariano Barbacid

Developmental Oncology Section, Basic Research Program, Frederick Cancer Research Facility, Frederick, Maryland 21701, USA

Induction of mammary carcinomas in rats by a single exposure to a carcinogen during sexual development often involves malignant activation of the Ha-ras-1 locus. Each of the Ha-ras-1 oncogenes present in tumours induced by N-nitroso-N-methylurea, but not in those induced by 7,12-dimethylbenz(a)anthracene, became activated by the same G→A transition, the type of mutation induced by N-nitroso-N-methylurea. These results are consistent with the notion that Ha-ras-1 oncogenes are directly activated by the carcinogen during initiation of neoplasia.

ELUCIDATION of the contribution of *ras* oncogene activation to the development of human neoplasia has been hampered by the unknown aetiology of most human cancers and the impossibility of experimental manipulations. Therefore, we have investigated animal tumour model systems that closely resemble human neoplasia. We have shown that transforming Ha-ras-1 genes can be reproducibly detected in mammary carcinomas induced in Buf/N rats by a single dose of N-nitroso-N-methylurea (NMU)¹. Similar results have been obtained in other

laboratories using other carcinogen-induced animal tumour systems²⁻⁴ as well as with chemically transformed cell lines⁵⁻⁷. These carcinogen-activated *ras* oncogenes have the same type of activating mutations as those present in human tumours^{1,8}, thus validating these animal tumour systems as adequate models for studying the role of *ras* oncogene activation in human neoplasia.

Understanding the role that *ras* genes play in the multi-step process of carcinogenesis requires precise definition of the stage

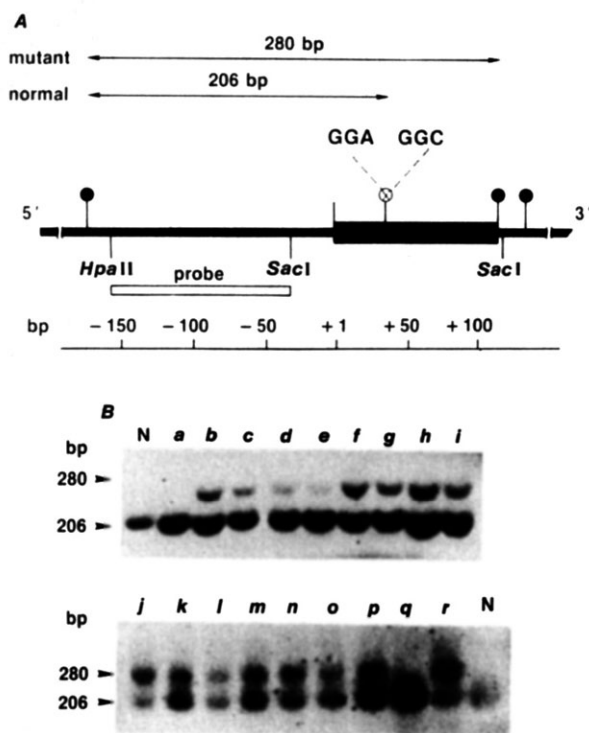


Fig. 1 **A**, Strategy for detection of *MnlI* RFLPs created by mutations in codon 12 or 13 of the rat *Ha-ras-1* locus. Wild-type *MnlI* restriction sites (●) near the first exon (solid box, nucleotides +1 to +111) are indicated. The position of the polymorphic *MnlI* site (GAGG, ⊗, nucleotides +35 to +38) relative to the sequences of the normal codon 12 (GGA) and codon 13 (GGC) is shown. The probe used is a 120-bp *HpaII* to *SacI* fragment (open box) comprising 5'-noncoding sequences derived from pNMU-1, a pBR322 derivative containing a transforming allele of the *Ha-ras-1* locus¹ and allowing specific detection of this locus. Probes encompassing exon-coding sequences (for example, the 154-bp *SacI* DNA fragment) detect an additional band, presumably derived from the *Ha-ras-2* pseudogene, that interferes with detection of the *MnlI* RFLPs. Arrows span the normal (206-bp) and polymorphic (280-bp) *MnlI* restriction fragments detected by the 120-bp *HpaII*/*SacI* probe. **B**, Use of *MnlI* RFLPs to detect *Ha-ras* oncogenes in DNA isolated from NMU-induced rat mammary carcinomas: 40 µg of DNA isolated from normal breasts (N) and NMU-induced mammary carcinomas (a-r) were digested for 16 h with 40 U *MnlI* using conditions recommended by the supplier (New England Biolabs). Digests were electrophoresed (18 h at 40 V) in horizontal 2.8% agarose gels (w/v) and blotted to either Zeta-Probe blotting membranes (Bio-Rad Laboratories) or nitrocellulose as described by Southern²³. Membranes were hybridized under stringent conditions (50% formamide, 5 × SSC, 42 °C) for 72 h to 10⁷ c.p.m. ml⁻¹ of ³²P-labelled probe obtained by nick-translation of the 120-bp *HpaII*/*SacI* DNA fragment described in **A**. Blots were exposed for 24 h to Kodak XAR-5 film at -70 °C in the presence of intensifier screens. Arrows indicate the positions of the normal (206-bp) and the polymorphic (280-bp) *MnlI* restriction fragments. Co-electrophoresed ΦX174 DNA digested with *HaeIII* served as size standards. Upper and lower panels represent two independent representative experiments.

at which they become activated. Recently, Balmain *et al.* have demonstrated the presence of transforming *ras* genes at the papilloma stage of 7,12-dimethylbenz(a)anthracene (DMBA)-induced skin carcinomas, suggesting that *ras* genes are activated early in tumour development⁹. Induction of rat mammary carcinomas by NMU does not proceed through identifiable pre-neoplastic stages¹⁰. We have taken advantage of two properties of this tumour system to investigate whether *Ha-ras-1* oncogenes are involved in initiation of carcinogenesis. First, NMU is known to specifically (≥99%) induce G → A transitions (see refs 11, 12 for reviews). Here, we demonstrate that each of the *Ha-ras-1*

Table 1 Malignant activation of the *Ha-ras-1* locus in NMU-induced rat mammary carcinomas*

Strain	Mammary carcinomas Tested	<i>Ha-ras-1</i> oncogene	Normal breasts Tested†	<i>Ha-ras-1</i> oncogene
Buf/N	38	31 (81%)	18	0
Sprague-Dawley	12	11 (92%)	8	0
Fischer 344	8	6 (75%)	ND	ND
Total	58	48 (83%)	26	0

* As determined by the *MnlI* RFLP assay. Correspondence between *MnlI* RFLPs and transforming activity was as high as 92% of tumour DNAs tested in more than one gene transfer assay. None of the normal breast DNAs was capable of inducing malignant transformation on transfection into NIH 3T3 cells.

† Includes four Buf/N and two Sprague-Dawley normal breasts obtained from tumour-bearing animals. ND, not done.

oncogenes present in NMU-induced mammary carcinomas became activated by G → A transitions in the second nucleotide of codon 12. No such mutations were observed in *Ha-ras-1* oncogenes present in mammary carcinomas induced by DMBA, a carcinogen that does not induce specific G → A transitions¹¹. These findings strongly implicate NMU as the agent directly responsible for the activation of *Ha-ras-1* oncogenes. Second, the highly labile nature of NMU under physiological conditions has defined initiation of carcinogenesis as occurring within hours after administration of the single tumour-inducing dose of this carcinogen. Therefore, if NMU is responsible for the G → A mutations found in each *Ha-ras* oncogene, the process of oncogene activation must be concomitant with, and presumably contributing to, initiation of carcinogenesis.

Mutagenesis in the *Ha-ras-1* locus

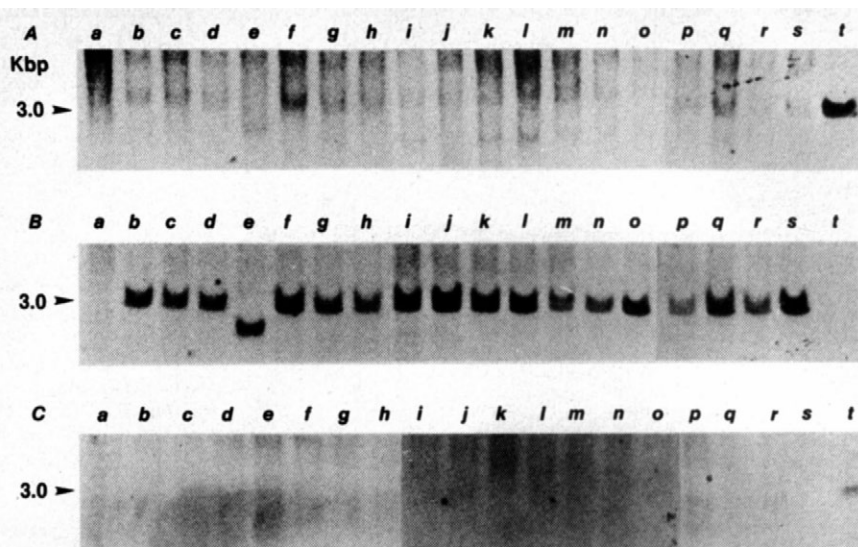
We have previously reported the presence of transforming *Ha-ras-1* genes in nine mammary tumours of inbred Buf/N rats induced by NMU¹. One of these oncogenes was molecularly cloned and its activating mutation identified as a G → A transition in the second nucleotide of codon 12 (ref. 1). This mutation eliminates a GAGG sequence that is specifically recognized by the restriction endonuclease *MnlI*, thus creating a diagnostic restriction fragment length polymorphism (RFLP). The first exon of the normal *Ha-ras-1* allele is located in two *MnlI* DNA fragments of 206 and 74 base pairs (bp) (Fig. 1A). Elimination of the *MnlI* cleavage site generates a single 280-bp *MnlI* DNA fragment. Wild-type (206 bp) and polymorphic (280 bp) DNA fragments were identified by Southern blot analysis using the 120-bp *HpaII*/*SacI* DNA fragment depicted in Fig. 1A as a probe.

Thirty-one of 38 DNAs isolated from NMU-induced tumours in Buf/N rats exhibited the diagnostic *MnlI* RFLP (Fig. 1B, Table 1). In addition to the 280-bp polymorphic fragment, each tumour DNA exhibited the normal 206-bp *MnlI* DNA fragment indicative of heterozygosity at this locus (Fig. 1B). Similar results were obtained with random-bred Sprague-Dawley (11 out of 12) and Fischer 344 (6 out of 8) rats, indicating the general validity of these observations. In total, 83% (48 out of 58) of all NMU-induced tumours tested scored as positive in the *MnlI*/RFLP assay. Of these 48 tumour DNAs, 36 (75%) induced malignant transformation of NIH 3T3 mouse cells in gene transfer experiments, verifying that transfection is a reliable assay for the detection of *ras* oncogenes.

None of 20 normal breast DNAs obtained from 14 Buf/N and 6 Sprague-Dawley rats exhibited the *MnlI* RFLP (Table 1). Moreover, the 10 tumour DNAs that did not exhibit the *MnlI* polymorphism failed to transform NIH 3T3 cells even when repeatedly tested in gene transfer experiments. These findings indicate that malignant activation of *Ha-ras-1* oncogenes in NMU-induced mammary carcinomas is the result of mutations specifically localized in a region no larger than four nucleotides.

Fig. 2 Use of synthetic oligonucleotide probes to determine the point mutations responsible for malignant activation of the *Ha-ras-1* locus in NMU-induced rat mammary carcinomas. **A**, Hybridization with probe Ha19-G³⁵ (5'-TGGGCGCTGGAGGCGTGGG-3'), which recognizes nucleotides +26 to +44 of the normal *Ha-ras-1* locus. **B**, Hybridization with probe Ha19-A³⁵ (5'-TGGGCGCTGAAGGCGTGGG-3'), which detects *Ha-ras-1* oncogenes activated by G→A transitions at nucleotide +35. **C**, Hybridization with probe Ha19-T³⁵ (5'-TGGGCGCTGTAGGCGTGGG-3'), which detects *Ha-ras-1* oncogenes activated by G→T transversions at nucleotide +35. **a**, DNAs isolated from NIH 3T3 cells; **b-s**, representative NIH 3T3 transformants derived from NMU-induced mammary carcinomas; **t**, NIH 3T3 cells co-transfected with the normal *Ha-ras-1* gene and pSV2-neo¹⁴ and selected for growth in the presence of G418. Arrowheads indicate the expected 3.0-kilobase (kb) *Hind*III DNA fragment of *Ha-ras-1* containing all coding sequences except the first 12 nucleotides¹. Abnormal migration is occasionally seen because the 3' *Hind*III cleavage site maps outside the *Ha-ras-1* locus and is occasionally lost during transfection. In control experiments, none of these DNAs hybridized to pBR322 or λ phage vector sequences. Co-electrophoresed DNA fragments of *Hind*III-cleaved λ c1857 DNA served as size standards. Results were compiled from two independent representative experiments.

Methods. 30 μ g of each DNA was digested with 240 U of *Hind*III using conditions recommended by the supplier (New England Biolabs). Digests were electrophoresed (12 h at 50 V) in horizontal agarose gels (0.7% w/v) which were subsequently prepared for hybridization as described previously by Kidd *et al.*²⁴. Dried gels were prehybridized for 1 h in hybridization buffer²⁴ containing 500 μ g ml⁻¹ of sonicated herring sperm DNA. Gels were subsequently hybridized for 18 h at 64 °C with individual oligonucleotide probes at a concentration of 5 ng ml⁻¹ (2.5×10^9 c.p.m. μ g⁻¹) in hybridization buffer containing 100 μ g ml⁻¹ sonicated herring sperm DNA. Hybridized gels were then washed with constant agitation as follows: four changes (300 ml) of 6 × SSC for 15 min each at room temperature, 400 ml of 6 × SSC for 30 min at room temperature; 500 ml of 6 × SSC at 66 °C was then poured onto the gel and allowed to cool to room temperature (30 min). Finally, the gel was rinsed with 200 ml of 6 × SSC at room temperature, dried and exposed to Kodak XAR-5 film for 18 h at -70 °C in the presence of intensifier screens. Labelling of oligonucleotide probes was performed under conditions allowing quantitative phosphorylation of 5' termini with polynucleotide kinase. Reactions (300 μ l) contained 0.01 nmol of individual nonadecamers, 1 nmol [γ -³²P]ATP (7,000 Ci mol⁻¹; NEN), 60 mM Tris-HCl (pH 7.8), 15 mM 2-mercaptoethanol, 10 mM MgCl₂ and 60 U T₄ polynucleotide kinase (New England Biolabs). Incubations were for 30 min at 37 °C and reactions were stopped by making mixtures 20 mM with respect to EDTA (pH 7.5). Labelled oligonucleotides were separated from unincorporated [γ -³²P]ATP by ion-exchange chromatography on Elutip-d columns (Schleicher & Schuell) according to the manufacturer's protocol. ³²P-oligonucleotides were eluted with 0.5 ml of 1 M NaCl, 20 mM Tris-HCl (pH 7.5), 1 mM EDTA and added directly to hybridization mixtures.



G³⁵→A mutations

We next examined whether the diagnostic *Mn*I RFLPs detected in each of the *Ha-ras-1* oncogenes present in NMU-induced mammary carcinomas were always the consequence of G→A mutations, which would implicate the mutagenic activity of NMU in the generation of these oncogenes. We used oligonucleotide probes to detect specific point mutations in genomic DNA¹³. We synthesized nonadecamers capable of identifying substitutions in position 35, the only nucleotide of the *Mn*I cleavage site that can alter the coding properties of the critical codon 12 of the *Ha-ras-1* gene. Oligomers included Ha19-G³⁵ (5'-TGGGCGCTGG*AGGCGTGGG-3', where the underlined nucleotides define the diagnostic *Mn*I cleavage site present in wild-type DNA and G* is the deoxyguanosine residue at position 35), Ha19-A³⁵ and Ha19-T³⁵. Ha19-A³⁵ and Ha19-T³⁵ possessed the same sequence as Ha19-G³⁵ except that the residue corresponding to position 35 (G* in Ha19-G³⁵) was substituted by either a deoxyadenosine in Ha19-A³⁵ or a thymidine in Ha19-T³⁵. The corresponding Ha19-C³⁵ oligonucleotide was not synthesized because G³⁵→C transversions would have created a polymorphic *Pst*I cleavage site that was not detected in any of the tumours tested (data not shown).

Oligonucleotide probes were hybridized to *Hind*III-cleaved DNAs isolated from representative NIH 3T3 transformants obtained from each of 36 NMU-induced mammary carcinomas that scored positive in gene transfer experiments. NIH 3T3 transformants were selected because they usually contain amplified oncogene sequences and lack the normal *Ha-ras-1* rat allele. Results of representative experiments are given in Fig. 1. At the discriminating temperature, each of the 36 NIH 3T3 transformants tested hybridized to Ha19-A³⁵ but not to Ha19-G³⁵ or Ha19-T³⁵, indicating that their transforming *Ha-ras-1* oncogenes carried identical G→A transitions in position 35. DNA isolated from NIH 3T3 cells containing a normal *Ha-ras-1* proto-oncogene (which was co-transfected with pSV2-neo¹⁴)

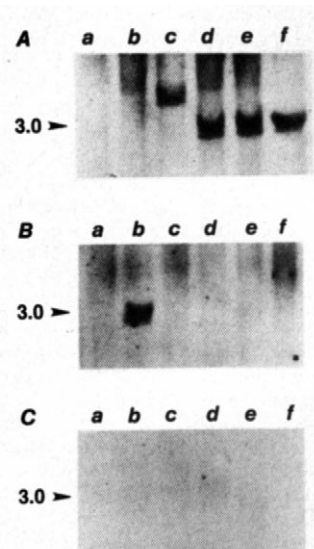


Fig. 3 Absence of point mutations in codon 12 of *Ha-ras-1* oncogenes present in mammary carcinomas induced by DMBA. **A**, Ha19-G³⁵ detects *Ha-ras-1* genes with a normal codon 12 (G at position +35); **B**, Ha19-A³⁵ detects G→A transitions in position +35; **C**, Ha19-T³⁵ detects G→T transversions in position +35. Arrowheads indicate the expected position of the 3.0-kb *Hind*III fragment as deduced from co-electrophoresed DNA fragments of *Hind*III-cleaved λ c1857 DNA. **a**, DNAs isolated from NIH 3T3 cells; **b**, NIH 3T3 transformant derived from an NMU-induced mammary carcinoma; **c-e**, representative NIH 3T3 transformants derived from DMBA-induced mammary carcinomas; **f**, NIH 3T3 cells co-transfected with the normal *Ha-ras-1* gene and pSV2-neo¹⁴ and selected for growth in the presence of G418. These DNAs were digested with *Hind*III and electrophoresed in 0.7% (w/v) agarose gels. Dried gels were hybridized to each of the three nonadecamer probes using the same conditions as described in Fig. 2 legend.

Table 2 Malignant activation of the Ha-ras-1 locus in NMU- and DMBA-induced rat mammary carcinomas

Carcinogen	Oncogene detection*	Ha-ras-1 oncogene	G ³⁵ → A ³⁵ mutation
NMU	48/58 (82%)	48/48 (100%)	36/36 (100%)
DMBA	3/14 (21%)	3/3 (100%)	0/3 (0%)

These results demonstrate that the same point mutation is responsible for malignant activation in NMU- but not in DMBA-induced rat mammary carcinomas.

* Results of NMU-induced tumours include *MnII* RFLP and gene transfer assays.

hybridized to Ha19-G³⁵ but not to Ha19-A³⁵ or Ha19-T³⁵ oligonucleotide probes (Fig. 2). These results demonstrate that each of the Ha-ras-1 oncogenes present in NMU-induced mammary carcinomas became activated by the same mutation, a G → A transition in the second nucleotide of the critical 12th codon.

Ha-ras-1 genes in DMBA-induced tumours

The striking specificity of our findings suggests that the activating G³⁵ → A mutations are the direct consequence of the mutagenic activity of NMU. However, we also considered the possibility that the presence of p21 proteins carrying glutamic acid residues in position 12 confers a selective growth advantage on the neoplastic mammary cells. To resolve this question we examined the mutations responsible for the activation of Ha-ras-1 oncogenes in mammary carcinomas induced by DMBA. DMBA, unlike NMU, forms large adducts with deoxyguanosine and deoxyadenosine residues that lead to excision repair, thus generating point mutations of undefined specificity¹¹. Representative NIH 3T3 transformants derived from each of the three DMBA-induced mammary carcinomas known to contain a transforming Ha-ras-1 oncogene efficiently hybridized to the Ha19-G³⁵ probe but not to the Ha19-A³⁵ or Ha19-T³⁵ probes (Fig. 3). These results indicate that DMBA-induced Ha-ras-1 oncogenes do not contain point mutations in either the 12th or adjacent codons. In agreement with these observations, none of these oncogenes exhibit the polymorphic *MnII* RFLP (data not shown).

Activation of Ha-ras-1 oncogenes in DMBA-induced tumours probably involves mutations in codon 61 (refs 15, 16). Although we have not defined the nature of such mutations, note that codon 61 of the rat Ha-ras-1 proto-oncogene is CAA^{1,17}. G → A transitions cannot occur in this codon because they will result in the generation of a terminator codon, TAA. Whatever the specific nature of DMBA-induced mutations, our results argue strongly against the possibility that the specific G³⁵ → A mutations of NMU-induced Ha-ras-1 oncogenes are the result of a positive growth selection process or a unique DNA repair system present in mammary cells. Instead, the results indicate that malignant activation of the Ha-ras-1 locus in NMU-induced mammary carcinomas must be the direct result of the mutagenic effect of NMU.

Biological implications

Induction of mammary carcinomas in rats by an injection of a single dose of NMU is one of the animal tumour systems most amenable to molecular analysis¹. Because of the specific mutagenic properties of NMU (G → A transitions) it has been possible to examine whether oncogenes can be the target of chemical carcinogens. Our results demonstrate that each of 36 NMU-induced Ha-ras-1 oncogenes carried the same activating G³⁵ → A mutation (Table 2). These findings strongly support the concept that these mutations are caused by NMU-directed methylation of the O⁶ position of the second deoxyguanosine (G³⁵) in the critical codon 12. If so, our observations represent the first identification of an oncogene as a target for a chemical carcinogen. Alternative explanations for the observed mutational specificity, such as specific G → A repair mechanisms in mammalian cells or frequent G → A substitutions during tumour development, are unlikely in view of our control studies with DMBA-induced Ha-ras-1 oncogenes (Table 2). Moreover, Guerrero *et al.* have recently shown that Ki-ras oncogenes reproducibly activated in mouse lymphomas induced by γ -radiation exhibit diverse activating mutations⁸ as expected from the type of DNA lesions induced by this carcinogen¹¹.

Our findings have further biological implications. Efficient induction of mammary carcinomas can be readily achieved by a single dose of NMU^{1,18}. Because of the highly labile nature of NMU, its mutagenic effect must occur within hours following its administration. Therefore, our results imply that malignant activation of the Ha-ras-1 locus by NMU-induced G → A mutations must be concomitant with the initiation of the carcinogenic process.

Several laboratories have shown that *ras* oncogenes cannot transform normal (primary) rodent embryo cells unless they acquire the capability of self-proliferation either spontaneously (that is, NIH 3T3 cells), by carcinogen treatment¹⁹ or by transfection of other oncogenes such as *c-myc*²⁰, adenovirus E1a²¹ or even Ha-ras²² when linked to strong transcriptional enhancers. It might be postulated that initiation of carcinogenesis by NMU requires the concomitant activation of two or more cooperating oncogenes, only one of which, Ha-ras-1, can be detected in gene transfer assays. But it is equally plausible that hormone-dependent proliferation, a normal physiological process required for tumour induction and (in some cases) progression, might provide conditions that are adequate for the phenotypic expression of the NMU-activated Ha-ras-1 oncogenes.

We thank Drs Steve Hughes, William Lijinsky and George Vande Woude for comments and for critically reading the manuscript. We also thank Dr P. Gullino for the DMBA-induced mammary tumours. H.Z. is the recipient of a Postdoctoral Fellowship from Fonds de la Recherche en Santé du Québec, Québec, Canada. Research sponsored by the National Cancer Institute, DHHS, under contract N01-CO-23909 with Litton Bionetics, Inc.

Note added in proof: We have now identified G³⁵ → A transitions in each of the Ha-ras-1 oncogenes present in the 48 NMU-induced mammary tumours used in this study. In addition, we have localized the activating mutation of the three DMBA-induced Ha-ras-1 oncogenes in the two deoxyadenosine residues of codon 61 (CAA).

Received 14 January; accepted 28 March 1985.

1. Sukumar, S., Notario, V., Martin-Zanca, D. & Barbacid, M. *Nature* **306**, 658-661 (1983).
2. Balmain, A. & Pragnell, I. B. *Nature* **303**, 72-74 (1983).
3. Eva, A. & Aaronson, S. A. *Science* **220**, 955-956 (1983).
4. Guerrero, I., Calzada, P., Mayer, A. & Pellicer, A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 202-205 (1984).
5. Padhy, L. C., Shih, C., Cowing, D., Finkelstein, R. & Weinberg, R. A. *Cell* **28**, 865-871 (1982).
6. Parada, L. F. & Weinberg, R. A. *Molec. cell. Biol.* **3**, 2298-2301 (1983).
7. Sukumar, S. *et al. Science* **223**, 1197-1199 (1984).
8. Guerrero, I., Villasante, A., Corcos, V. & Pellicer, A. *Science* **225**, 1159-1162 (1984).
9. Balmain, A., Ramsden, M., Bowden, G. T. & Smith, J. *Nature* **307**, 658-660 (1984).
10. Gullino, P. M., Pettigrew, H. M. & Grantham, F. H. *J. natn. Cancer Inst.* **54**, 401-409 (1975).
11. Singer, B. & Kusmierek, J. T. *A. Rev. Biochem.* **51**, 655-693 (1982).

12. Pegg, A. E. *Cancer Invest.* **2**, 223-231 (1984).
13. Conner, B. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 278-282 (1983).
14. Southern, P. J. & Berg, P. *J. molec. appl. Genet.* **1**, 327-341 (1982).
15. Yuasa, Y. *et al. Nature* **303**, 775-779 (1983).
16. Taparowsky, E., Shimizu, K., Goldfarb, M. & Wigler, M. *Cell* **34**, 581-586 (1983).
17. Dhar, R. *et al. Science* **217**, 934-937 (1982).
18. McCormick, D. L., Adamowski, C. B., Fiks, A. & Moon, R. C. *Cancer Res.* **41**, 1690-1694 (1981).
19. Newbold, R. F. & Overell, R. W. *Nature* **304**, 648-651 (1983).
20. Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596-602 (1983).
21. Rulley, H. E. *Nature* **304**, 602-606 (1983).
22. Spandidos, D. A. & Wilkie, N. M. *Nature* **310**, 469-475 (1984).
23. Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
24. Kidd, V. J., Wallace, R. B., Itakura, K. & Woo, S. L. C. *Nature* **304**, 230-234 (1983).

Galactic mergers, starburst galaxies, quasar activity and massive binary black holes

C. Martin Gaskell

Astronomy Department, University of Texas, Austin, Texas 78712, USA

Many quasar-like objects show evidence for massive binary black holes¹. The recent discovery² of a massive ($5 \times 10^6 M_\odot$) object in the centre of the local group dwarf elliptical M 32 greatly raises the probability of forming such binaries through galactic mergers. Here I argue that the enhancement of all kinds of activity (quasar-like activity and star formation) in galaxies with companions is not so much a consequence of tidal interaction between the massive galaxies as the result of collisions with their dwarf satellites.

Recent studies have shown that interactions between galaxies play a role in stimulating nuclear activity³. Imaging studies of high luminosity quasars^{4,5} reveal that more than 30% have normal companion galaxies very close to them (here I use the term 'normal' galaxy to refer to any galaxy with a mass of $10^{11} M_\odot$ or more and refer to companion galaxies with masses much less than this as 'satellite' galaxies). Surveys of normal galaxies show that close pairs of galaxies are more likely to show activity characterised by optical emission lines^{6,7} and radio emission⁸ than are isolated galaxies. These results support the idea that most or all normal sized galaxies have massive black holes in their centres³. For simplicity I will assume that all normal galaxies have massive ($\geq 10^6 M_\odot$) black holes in their centres (it makes little difference here if the percentage is actually as low as 20%, as could be the case). The enhanced nuclear activity in pairs of galaxies would seem also to support an idea proposed by Toomre and Toomre⁹ that close encounters may drive gas into the nuclear regions of a galaxy and either feed the central black hole directly or lead to a burst of star formation within the inner nuclear disc⁷ which could indirectly increase the fuel supply rate. In addition to the enhanced nuclear activity close pairs of galaxies have enhanced star formation rates (as evidenced by far infrared emission¹⁰).

There are two problems with the Toomre and Toomre scenario. First, although a close pair of galaxies (with a projected distance of ~ 3 Holmberg diameters or less) is more likely to be active (as a Seyfert or radio galaxy), the majority of such galaxies (>75%) do not show nuclear activity. The mechanism is therefore less than 100% efficient or there is some other factor involved. The second problem is that although close encounters can wreak havoc on the outer parts of galaxies, the motions of the stars and gas in the tightly gravitationally bound cores are unaffected.

The solution I propose to these problems is that it is not actually the detected neighboring galaxy that is triggering quasar activity (or star formation) but instead a direct collision with one of the several dwarf satellite galaxies in the system. There is then no need to invoke some subtle action at a distance (tidal?) effect. Although the statistics of these dwarf satellite galaxies are poorly known, since they are difficult to detect beyond our local group, there are probably several (2–10) associated with each normal galaxy like our own¹¹. I will here use the term 'dwarf' galaxy to refer to ellipticals and irregulars with absolute magnitudes in the approximate range $-14 > M_V > -18$ (and hence ignore systems such as the Draco dwarf spheroidal). In our own local group there are two magellanic irregulars associated with our galaxy and four dwarf ellipticals associated with M31 in Andromeda.

In the scenario I propose here the statistical enhancement of quasar-like activity in close pairs of galaxies arises because of the increased probability of a merger with an unseen dwarf satellite galaxy when two normal galaxies are close. A 10^8 – $10^{10} M_\odot$ dwarf galaxy near a Seyfert is hard to detect with present techniques and the only sign of a collision or merger might be the enhanced quasar-like activity and/or a burst of star formation (perhaps creating a 'starburst' galaxy). I suggest that the type of activity produced is a function of the Hubble type of the colliding satellite galaxy and the impact parameter. The amount of visible damage is also obviously going to depend on the mass of the dwarf. If a magellanic irregular galaxy hits a spiral galaxy anywhere the most probable result is a burst of star formation. If it hits the nucleus the resulting gas input and star formation could result in quasar-like activity as well. On the other hand the impact of a small elliptical galaxy, devoid of dust and gas, on the outer regions of a spiral galaxy will probably be much less spectacular. Perhaps only a distortion of the disc and spiral arms of the galaxy will occur (a dwarf elliptical satellite has typically 1% of the mass of a normal spiral and that the stars themselves do not collide).

The collision of a dwarf elliptical with a nucleus containing a dormant (or almost dormant) black hole will be more interesting. Unlike magellanic irregulars, the majority of dwarf ellipticals have strong central mass concentrations¹². Tonry has produced powerful evidence that the 'spike' in the centre of M 32 is caused by a $5 \times 10^6 M_\odot$ black hole. It would seem likely that the spikes in the luminosity profiles of other dwarf ellipticals are due to similar black holes. In a collision, if the impact parameter is small enough, the nucleus of the dwarf elliptical will be captured by the nucleus of the larger galaxy. The two black holes will go into orbit^{1,13}. The distribution of stars in phase-space around each black hole will be severely perturbed, resulting in an increased fuelling of one or both black holes. I have proposed elsewhere¹ that the displaced broad-line peaks seen in many objects are the result of such orbital motion and argued that the velocities seen are consistent with the evolution of supermassive binary black holes. The discovery of a black hole in a dwarf elliptical now makes the creation of such binaries much more probable since there are an order of magnitude more dwarf galaxies than normal galaxies and the probability of mergers is correspondingly greater.

Detailed computer simulations of mergers of galaxies of all Hubble types and sizes taking into account both stars and gas need to be undertaken, as collisions of normal galaxies with dwarf satellite galaxies offer a likely explanation of both starburst galaxies and quasar-like activity. The Hubble Space Telescope could well reveal the results of the collisions when high spatial resolution studies of active galaxies are made.

I am grateful to Antoinette de Vaucouleurs, Gérard de Vaucouleurs and Derek Wills for useful comments. Partial financial support from McDonald Observatory is gratefully acknowledged.

Received 15 October 1984; accepted 1 February 1985.

- Gaskell, C. M. *Liège Astrophysical Colloq.* **24**, 473–477 (1983).
- Tonry, J. L. *Astrophys. J. Lett.* **283**, L27–L30 (1984).
- Gaskell, C. M. *Nature* **310**, 102 (1984).
- Stockton, A. N. *Astrophys. J.* **257**, 33–39 (1982).
- Hutchings, J. B. & Campbell, B. *Nature* **303**, 584–586 (1983).
- Stauffer, J. R. *Astrophys. J.* **262**, 66–80 (1982).
- Kennicutt, R. C. & Keel, W. C. *Astrophys. J. Lett.* **279**, L5–L9 (1984).
- Hummel, E. *Astr. Astrophys.* **96**, 111–119 (1981).
- Toomre, A. & Toomre, J. *Astrophys. J.* **178**, 623–666 (1972).
- Joseph, R. D., Meikle, W. P. S., Robertson, N. A. & Wright, G. S. *Mon. Not. R. astr. Soc.* **209**, 111–122 (1984).
- Reaves, G. *Astrophys. J. Suppl.* **53**, 375–395 (1983).
- Reaves, G. in *The Evolution of Galaxies and Stellar Populations* (eds Tinsley, B. M. & Larson, R. B.) 39–41 (Yale University Observatory, New Haven).
- Begelman, M. C., Blandford, R. D. & Rees, M. J. *Nature* **287**, 307–309 (1980).

Evidence for a 150–200-km thick Archaean lithosphere from diamond inclusion thermobarometry

F. R. Boyd*, J. J. Gurney† & S. H. Richardson‡

* Geophysical Laboratory, Carnegie Institution of Washington, Washington DC 20008, USA

† Department of Geochemistry, University of Cape Town, Rondebosch, South Africa

‡ Laboratoire de Geochemie et Cosmochimie, Institut de Physique du Globe, 4 Place Jussieu, 75230 Paris Cedex 05, France

Occurrences of diamondiferous kimberlites in southern Africa are concentrated within the boundaries of the Kaapvaal craton where crustal rocks are as old as 3,500 Myr (ref. 1). Garnet inclusions from diamonds that erupted in the Finsch kimberlite and kimberlites of the Kimberley group, situated in the southern part of the craton, have Rb/Sr and Sm/Nd model ages of 3,200–3,300 Myr (ref. 2). These kimberlites were erupted ~100 Myr ago^{3,4}; thus the diamonds are xenocrysts. The localization of diamond-bearing kimberlites in the craton and the eruption of Archaean diamonds in late Cretaceous time are evidence that these diamonds crystallized and were immobilized in a root extending to depths of 150–200 km. Our applications of thermobarometry to silicate inclusions in the diamonds suggest that ambient temperatures at those depths 3,000 Myr ago were 900–1,200 °C, similar to the range predicted from present-day heat flow⁵.

Monomineralic inclusions of olivine, garnet and pyroxenes are common in diamonds, and separate inclusions of two minerals in a single diamond have been found in many specimens⁶. Equilibration temperatures for these inclusion sets can be estimated by application of the same methods of geothermometry that have been widely used for garnet peridotite xenoliths^{7–9}, but it is important to recognize a marked difference in the circumstances of equilibration of the xenoliths and the inclusions.

Mineral grains in mantle rocks erupted as xenoliths were in contact with each other (and possibly with pore fluids) and had the potential for chemical exchange at high temperature until the time of eruption. Mineral grains that are included and isolated in diamonds, however, have chemical compositions that were fixed at the time of crystallization of the diamonds. The isolation of diamond inclusions is manifest in the marked differences in Nd and Sr isotope ratios obtained for garnet inclusions in diamonds and similar garnets from disaggregated mantle rocks found in kimberlite concentrates².

A crystallization history that is different from that of the inclusions has been experienced by three garnet lherzolite xenoliths from southern Africa that have diamond intergrown with primary silicate grains^{10,11}. These and numerous non-diamondiferous, low-temperature garnet peridotites from many kimberlite pipes have estimated temperatures and depths of origin that approximate a continental shield geotherm^{7,12}. The blocking temperatures for exchange equilibria used for peridotite thermometry have been estimated to be 700 °C (ref. 13) and 900 °C (ref. 14) on the basis of observed compositional profiles and experimental diffusion data. All three diamond-bearing lherzolites have equilibration temperatures in the range 1,050–1,150 °C (refs 7, 10). The least favourable estimate for the rate of equilibration at such temperatures¹⁴ makes it probable that these peridotites reached chemical equilibrium within 10–100 Myr of the time of eruption. Thus, the times at which compositional relations were established for Archaean diamond inclusions and for the diamondiferous peridotite xenoliths erupted 100 Myr ago are believed to differ by ~3,000 Myr.

The assumption that mineral grains forming separate inclusions in a single diamond were in chemical equilibrium at the time of crystallization is supported by the fact that multiple inclusions of one mineral have the same composition¹⁵. The

relative homogeneity of xenolith minerals¹⁶ and the fact that coexisting olivine, garnet and pyroxene have systematic partition relations for most major and minor elements⁷ are strong evidence of chemical equilibration for these rocks.

Diamond crystallization temperatures provide an upper limit for ambient temperatures at depth in the mantle at the times of diamond crystallization. The nature of the crystallization events in which natural diamonds have formed is not known; they may have been igneous or metasomatic or both. Nevertheless, these events can only have raised surrounding temperatures to values that were locally above ambient. The lower range of crystallization temperatures estimated for diamond inclusions is thus an approximation for ambient temperatures.

Comparative evaluation of element exchange reactions that can be used as thermometers for garnet-bearing, Mg-rich mineral assemblages has shown that application of the diopside–solvus method yields superior accuracy and precision⁷. The solvus cannot, however, be used for eclogitic assemblages or for diopside-free rocks and diamond inclusion sets. With such assemblages, comparable results can be obtained with Mg/Fe partition reactions, although in application to peridotites these Mg/Fe thermometers produce a scatter in estimated equilibration temperatures that is probably caused by the presence of small but unknown amounts of iron (III) (ref. 7).

Two Mg/Fe partition relations that have wide applicability in the study of diamond inclusions are clinopyroxene–garnet¹⁷ and olivine–garnet¹⁸. The effects on both of pressure and of the non-ideal substitution of Ca in garnet have been evaluated. The precision that can be obtained with the olivine–garnet thermometer has been estimated to be ± 50 °C for favourable conditions of temperature ($<1,300$ °C) and composition¹⁸, and the uncertainty associated with the garnet–diopside experimental data is a comparable value of $\pm 5\%$ of the temperature¹⁷.

Peridotitic garnets included in diamonds are commonly Cr-rich with Cr₂O₃ contents ranging up to 15 wt%. The effects of this solid solution on the garnet–olivine (Mg/Fe) thermometer have not yet been evaluated. Garnets in low-temperature lherzolite xenoliths have an overlapping range for Cr₂O₃ of 2–9 wt%, and for most of these rocks, temperatures estimated with the garnet–olivine thermometer agree with those based on the diopside solvus within 100 °C (ref. 19). There is no discernible relationship between estimated equilibration temperature and Cr content of garnet for the peridotite and diamond-inclusion assemblages. Hence, an effect of Cr solid solution on the garnet–olivine thermometer seems unlikely to be important.

Analyses for 31 olivine–garnet inclusion pairs from Finsch diamonds show that 29 of the garnets are low-Ca pyrope¹⁵, typical of the most common peridotitic garnets found included in diamonds. These are different specimens but are similar in composition to the garnet inclusions for which Archaean ages were obtained². Crystallization temperatures estimated for these Finsch olivine–garnet inclusion sets have the range 900–1,300 °C at an assumed pressure of 50 kbar, within the diamond stability field¹⁹. Only two of the pairs, however, have crystallization temperatures above 1,150 °C. Most, with temperatures of 900–1,150 °C, are believed to best reflect crystallization in near ambient conditions.

The olivine–garnet (Fe/Mg) thermometer has a small pressure sensitivity, ~ 5 °C kbar⁻¹. Estimates for the crystallization temperatures of the 29 lower-temperature inclusion pairs thus plot in a pressure–temperature (P – T) band 250 °C wide with a slope of 5 °C kbar⁻¹ lying in the diamond stability field (Fig. 1). One diamond inclusion set from Finsch contains the assemblage enstatite–olivine–garnet²⁰, and estimation of equilibration pressure using the Al content of the enstatite, as well as temperature using the olivine–garnet thermometer, can be made for this assemblage. A point for the three-phase set plots in the diamond stability field and within the P – T band established for the olivine–garnet pairs. Estimates for the crystallization temperatures and depths of these inclusions plot close to a shield geotherm calculated for a heat flow of 40 mW m⁻² (ref. 5).

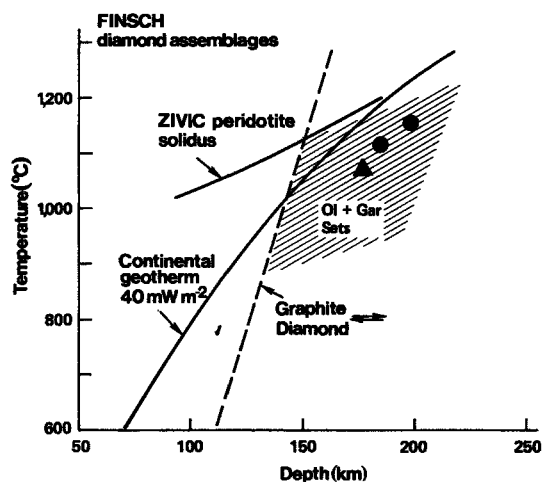


Fig. 1 Estimated equilibration conditions for diamondiferous assemblages from the Finsch kimberlite compared with the diamond-graphite equilibrium curve³⁵, a continental geotherm calculated for a heat flow of 40 mW m^{-2} (ref. 5), and the solidus for peridotite in the presence of H_2O and CO_2 (ZIVIC)²⁹. The band defined by ruled lines is for temperature estimates made by the (Fe/Mg) olivine-garnet method¹⁸, for 29 pairs of olivine and garnet inclusions. $\blacktriangle$, Olivine-garnet-orthopyroxene assemblage in a single diamond²⁰ for which temperature was estimated by the (Fe/Mg) olivine-garnet thermometer and pressure was obtained by the Al-enstatite method^{7,36}. $\bullet$, Diamondiferous lherzolite xenoliths¹⁰; temperatures for these were estimated from the diopside solvus³⁷, and pressures, from the alumina content of enstatite^{7,36}.

Equilibration conditions estimated for two diamond-bearing garnet lherzolite xenoliths from the Finsch pipe also plot in the diamond stability field in proximity to the inclusion data (Fig. 1). The fact that the equilibration conditions for most of the inclusions and the lherzolites are similar and concordant with the shield geotherm suggests that these P - T conditions were near ambient and that they have changed little in $\sim 3,000$ Myr.

Whether there are age differences either for diamonds of eclogitic and peridotitic affinities or related to their location in the Kaapvaal craton is not yet established, but all data available indicate Precambrian ages for diamond crystallization. Diamonds with garnet inclusions from the Finsch and Kimberley pipes dated at 3,200–3,300 Myr (ref. 2) have a peridotite paragenesis. Diamonds with both eclogitic and peridotitic affinities are present in the Premier kimberlite, and sulphide inclusions in Premier diamonds have yielded a model Pb age of 1,200 Myr (ref. 21); the paragenesis of the dated specimens, however, is unknown. Isotope analyses for Pb in clinopyroxenes from eclogite xenoliths from the Roberts Victor kimberlite define an apparent age of 2,500 Myr (ref. 21), whereas Sm/Nd and Rb/Sr determinations for other Roberts Victor samples indicate a model age that may be as much as 4,200 Myr (ref. 22). Eclogite xenoliths studied in these investigations do not contain diamonds, but many other Roberts Victor eclogites are diamondiferous. Ancient, alluvial diamonds of unknown paragenesis and primary source occur in the 2,500-Myr-old Witwatersrand conglomerates in the central part of the Kaapvaal craton²³. High $^3\text{He}/^4\text{He}$ values found in some South African diamonds of unknown specific source suggest diamond crystallization in the early Archaean²⁴. A model argon age of 3,100 Myr has been obtained for a single diamond from Arkansas²⁵, but similar measurements have not yet been performed on South African diamonds.

Crystallization temperatures estimated for diamonds from seven Kaapvaal kimberlites show the large range of 900–1,400 °C (Fig. 2). There is a difference in the estimated temperatures for diamonds with eclogitic and peridotitic inclusions; temperatures for the eclogitic suite are confined to the range 1,100–1,400 °C,

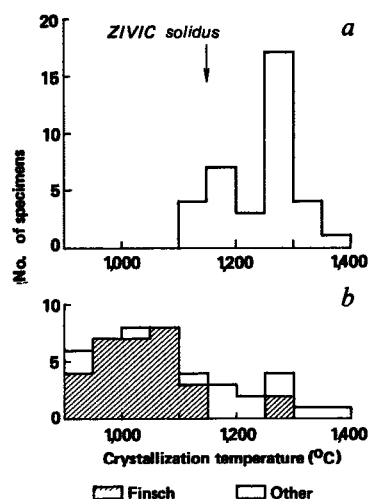


Fig. 2 Equilibration temperatures for sets of diamond inclusions of both eclogitic (a) and peridotitic (b) paragenesis estimated for a pressure of 50 kbar (160 km). Eclogitic inclusions are from diamonds erupted in the Roberts Victor, Orapa and Premier kimberlites. Peridotitic inclusions are from diamonds originating at Finsch, Jagersfontein, Koffiefontein, Roberts Victor, Premier and an unknown locality. Temperatures for eclogitic inclusions were estimated with the garnet-clinopyroxene (Fe/Mg) thermometer¹⁷. Those for peridotitic inclusions were estimated primarily with the olivine-garnet (Fe/Mg) relation¹⁸, but the diopside solvus³⁷ and garnet-clinopyroxene (Fe/Mg) thermometer¹⁷ were also used.

whereas the peridotitic suite has the full range 900–1,400 °C, but with a concentration in the range 900–1,100 °C. Geographical differences, however, appear in these data. Most of the inclusion sets with crystallization temperatures $< 1,100$ °C are from the Finsch pipe (Fig. 2). Other suites of diamonds with inclusions of both peridotitic and eclogitic parageneses (in separate diamonds) from the Premier and Roberts Victor kimberlites have higher crystallization temperatures, predominantly $> 1,200$ °C (refs 26, 27). Differences in mean carbon isotope composition between diamonds from the Finsch and Premier kimberlites have also been noted²⁸.

The beginning of melting for peridotite (or kimberlite) in the presence of H_2O and CO_2 at 50 kbar is $\sim 1,150$ °C (ref. 29). This fact led Boyd and Finnerty¹⁹ to suggest that peridotitic diamonds with crystallization temperatures below 1,150 °C might be of sub-solidus origin. An obvious corollary to this suggestion is evident from the data in Fig. 2; that is, eclogitic diamonds and perhaps some peridotitic diamonds with crystallization temperatures $> 1,150$ °C might be igneous. Diamonds with crystallization temperatures in the lower part of the range shown in Fig. 2 are believed to approach ambient conditions most nearly, regardless of the nature of the fluid from which they were precipitated.

Evidence from many xenolith suites supports the concept of a peridotite-rich, cratonic lithosphere with a thickness of 150–200 km. Recent seismic studies have provided data that are concordant with this model³⁰, and evidence for seismic velocity drops in the depth range 150–200 km (ref. 31) may correlate with geotherm inflections. Jordan³² has suggested that thick continental tectospheres developed as early as 2,000–2,800 Myr.

Burke and Kidd³³ have argued that the rarity of minimum-melting granites in the Superior Province is evidence that Archaean gradients beneath shields were not much steeper than are those of the present time. Analogous evidence led Davies³⁴ to propose that Archaean cratons were underlain by relatively cool 'root zones' or lithosphere that extended to a depth of 200 km where the palaeotemperature was near 1,500 °C.

The results obtained from the chemical and isotopic studies of diamond inclusions are evidence that the Kaapvaal craton, and perhaps other cratons, were stabilized earlier (3,300 Myr) than hitherto believed. Moreover, these data appear most concordant with thermal models for the Earth during the Archaean in which increased heat generation from radioactive sources was lost primarily through the ocean basins (see ref. 34).

Received 16 October 1984; accepted 27 March 1985.

- Hamilton, P. J., Evenson, N. M., O'Nions, R. K., Smith, H. S. & Erlank, A. J. *Nature* **279**, 298-300 (1979).
- Richardson, S. H., Gurney, J. J., Erlank, A. J. & Harris, J. W. *Nature* **310**, 198-202 (1984).
- Davis, G. L. in *Extended Abstr. 2nd int. Kimberlite Conf., Santa Fe* (American Geophysical Union, Washington, DC, 1977).
- Smith, C. B. *Nature* **304**, 51-54 (1983).
- Pollack, H. N. & Chapman, D. S. *Tectonophysics* **38**, 279-296 (1977).
- Meyer, H. O. A. & Tsai, H.-M. *Miner. Sci. Engng* **8**, 242-261 (1976).
- Finnerty, A. A. & Boyd, F. R. *Geochim. cosmochim. Acta* **48**, 15-27 (1984).
- Carswell, D. A. & Gibb, F. G. F. *Contr. Miner. Petrol.* **74**, 403-416 (1980).
- Boyd, F. R. *Geochim. cosmochim. Acta* **37**, 2533-2546 (1973).
- Shee, S. R., Gurney, J. J. & Robinson, D. N. *Contr. Miner. Petrol.* **81**, 79-87 (1982).
- Dawson, J. B. & Smith, J. V. *Nature* **254**, 580-581 (1975).
- Boyd, F. R. *Geology* **12**, 528-530 (1984).
- Smith, D. & Wilson, C. R. *Am. Miner.* **70**, 30-39 (1985).
- Harte, B. & Freer, R. *Terra Cognita* **2**, 273 (1982).
- Gurney, J. J., Harris, J. W. & Rickard, R. S. in *Proc. 2nd int. Kimberlite Conf. Vol. 1* (eds Boyd, F. R. & Meyer, H. O. A.) 1-15 (American Geophysical Union, Washington, DC, 1979).
- Boyd, F. R. & Finger, L. W. *Yb. Carnegie Inst. Wash.* **74**, 519-525 (1975).
- Ellis, D. J. & Green, D. H. *Contr. Miner. Petrol.* **71**, 13-22 (1979).
- O'Neill, H. St. C. & Wood, B. J. *Contr. Miner. Petrol.* **70**, 59-70 (1979).
- Boyd, F. R. & Finnerty, A. A. *J. geophys. Res.* **85**, 6911-6918 (1980).
- Tsai, H.-M., Meyer, H. O. A., Moreau, J. & Milledge, H. J. in *Proc. 2nd int. Kimberlite Conf. Vol. 2* (eds Boyd, F. R. & Meyer, H. O. A.) 16-26 (American Geophysical Union, Washington, DC, 1979).
- Kramers, J. D. *Earth planet. Sci. Lett.* **42**, 58-70 (1979).
- Jagoutz, E., Spettel, B. & Wänke, H. *Meteoritics* **18**, 319-320 (1983).
- Williams, A. F., *The Genesis of the Diamond* Vol. 2 (Ernest Benn, London, 1932).
- Ozima, M. & Zashu, S. *Science* **219**, 1967-1968 (1983).
- Melton, C. E. & Giardini, A. A. *Geophys. Res. Lett.* **7**, 461-464 (1980).
- Gurney, J. J., Harris, J. W. & Rickard, R. S. (in preparation).
- Gurney, J. J., Harris, J. W. & Rickard, R. S. in *Kimberlites: The Mantle and Crust-Mantle Relationships* (ed. Kornprobst, J.) 25-32 (Elsevier, Amsterdam, 1984).
- Deines, P., Gurney, J. J. & Harris, J. W. *Geochim. cosmochim. Acta* **48**, 325-342 (1984).
- Eggler, D. H. & Wendlandt, R. F. in *Proc. 2nd int. Kimberlite Conf. Vol. 1* (eds Boyd, F. R. & Meyer, H. O. A.) 330-338 (American Geophysical Union, Washington, DC, 1979).
- Anderson, D. L. & Bass, J. D. *Geophys. Res. Lett.* **11**, 637-640 (1984).
- Given, J. W. & Helmberger, D. V. *J. geophys. Res.* **85**, 7183-7194 (1980).
- Jordan, T. H. *Nature* **274**, 544-548 (1978).
- Burke, K. & Kidd, W. S. F. *Nature* **272**, 240-241 (1978).
- Davies, G. F. *Earth planet. Sci. Lett.* **44**, 231-238 (1979).
- Kennedy, C. S. & Kennedy, G. C. *J. geophys. Res.* **81**, 2467-2470 (1976).
- MacGregor, I. D. *Am. Miner.* **59**, 110-119 (1974).
- Lindsley, D. H. & Dixon, S. S. *Am. J. Sci.* **276**, 1285-1301 (1976).

Geochemistry of Holocene volcanic rocks associated with ridge subduction in Baja California, Mexico

G. Rogers*, A. D. Saunders*, D. J. Terrell†, S. P. Verma‡ & G. F. Marriner§

* Department of Geology, University of Leicester, Leicester LE1 7RH, UK

† Instituto de Geofísica, Universidad Nacional Autónoma de México, Ciudad Universitaria, 04510 Mexico DF, Mexico

‡ Departamento de Geotermia, Instituto de Investigaciones Electricas, Apartado Postal 475, Cuernavaca, Mor. 6200, Mexico

§ Department of Geology, Bedford College, Regent's Park, London NW1 4NS, UK

Although the geochemistry of magmatism associated with the subduction of oceanic lithosphere is well understood¹⁻⁶, the geochemical signature of lavas produced shortly after subduction of an oceanic spreading centre has not been characterized. The Baja California peninsula in Mexico provides an ideal scenario to study the latter process. Geophysical models indicate that, in response to oblique collision of the ancestral East Pacific Rise (EPR) with the western seaboard of North America at ~29 Myr, subduction of the Farallon plate was succeeded by the progressive southwards development of a transform fault system⁷⁻⁹. At 12.5 Myr a considerable length of the EPR was simultaneously subducted beneath Baja, thus terminating subduction processes along this segment¹⁰. By 3.5 Myr the locus of transcurrent faulting

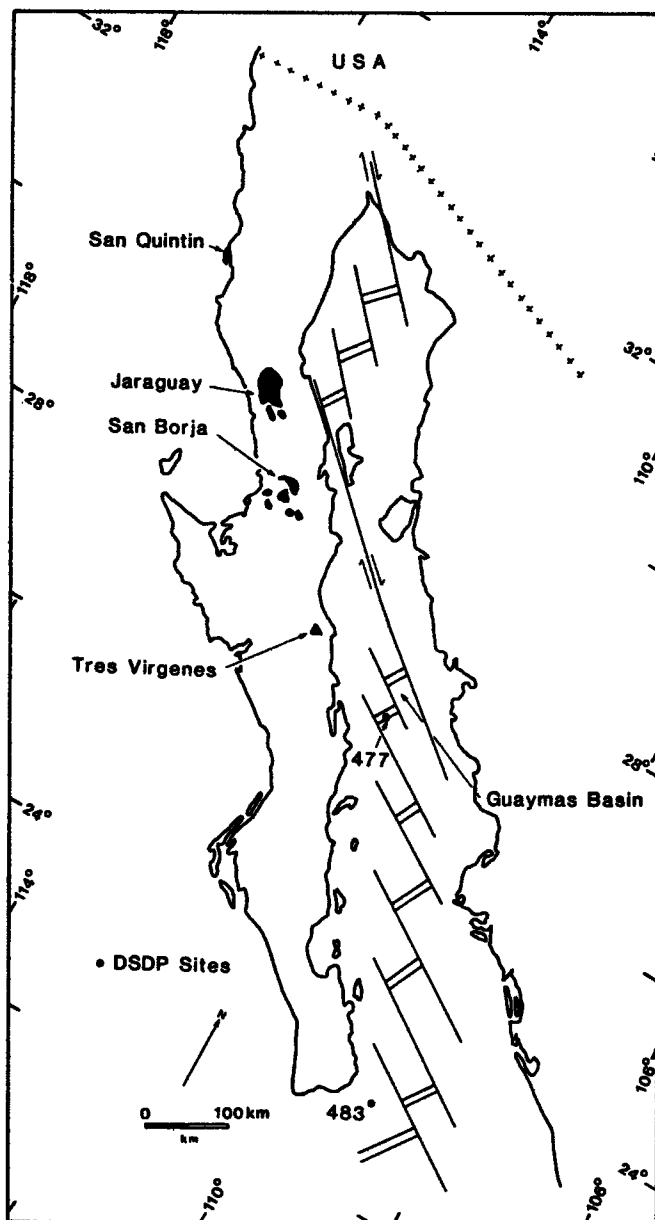


Fig. 1 Map depicting areas of Upper Cenozoic-Holocene volcanism in Baja California showing the three volcanic fields of Baja California Norte, and the Tres Virgenes volcano of Baja California Sur.

had switched from the west of Baja to within the developing Gulf of California, coupling Baja to the Pacific plate¹¹⁻¹³. We present here data for Holocene volcanic rocks from Baja California Norte which show that post-ridge subduction volcanism has a distinctive geochemistry. These results have important consequences for the interpretation of ancient orogenic belts.

In sympathy with the changing tectonic regimes, there are variations in the geochemistry of associated volcanism. Calc-alkaline lavas related to subduction were widespread until ~15 Myr (refs 11, 16, 17) since when there has been an increase in tholeiitic and alkaline products^{11, 14-18}, although the currently active Tres Virgenes stratovolcano¹⁹ is also calc-alkaline²⁰. As shown in Fig. 1, Holocene volcanism in northern Baja California is confined to three main areas, namely around San Quintin, Jaraguay and San Borja, although there are isolated flows elsewhere in the region^{16, 17, 21, 22}. Additionally, the lavas at San Quintin contain the only known occurrences on the peninsula of ilherzolite nodules²³.

All the rocks studied form blocky flows erupted from small cones, and are fresh and sparsely porphyritic (<5%). They are morphologically very young: flows at San Quintin which overlie

Table 1 Representative analyses of Holocene lavas from Baja California Norte, the Gulf of California and Isla Cook, Chile

Sample	San Quintin		Jaraguay		San Borja	Guaymas	Gulf of	Isla
	SQ1.1	SQ9.1	J14.1	J14.3	SB19.1	Basin* 477A-2-2	California† 483-24-1	Cook‡ 4-151
SiO ₂	46.26	47.17	54.64	56.30	55.92	49.00	48.40	57.65
TiO ₂	2.93	2.19	1.09	1.18	2.15	1.76	2.41	1.23
Al ₂ O ₃	15.74	15.32	15.86	16.28	15.59	17.40	14.53	17.22
Fe ₂ O ₃	2.07	1.73	0.86	0.89	0.95	9.66§	13.67§	2.03¶
FeO	10.33	8.64	4.31	4.45	4.76	—	—	2.57¶
MnO	0.20	0.17	0.08	0.08	0.08	0.15	0.22	0.08
MgO	6.79	9.04	6.19	6.20	5.00	7.23	6.60	4.60
CaO	8.27	9.26	8.10	8.00	8.13	11.81	11.50	8.79
Na ₂ O	3.70	3.19	4.31	4.26	4.29	3.10	2.64	4.23
K ₂ O	2.02	1.47	1.83	1.84	2.26	0.25	0.12	0.43
P ₂ O ₅	0.80	0.51	0.53	0.53	0.85	0.23	0.20	0.35
LOI	-0.05	-0.23	0.68	0.49	0.46	—	0.06	0.51
Total	99.06	98.46	98.48	100.50	100.44	100.59	100.35	99.69
Sr	758	565	2,123	2,068	2,609	241	109	2,045
Rb	37	28	7.7	8.9	9.3	1	<1	1
Ba	470	377	812	804	1,050	22	37	132
Zr	301	203	158	166	257	123	144	195
Nb	60	41	4.7	4.9	10.4	4	4	14
Y	36	26	8.3	7.9	11.4	29	49	11
Cr	87	277	209	205	84	191	140	114
Ni	86	162	157	150	85	66	61	60
La	43	31	30	30	45	5.5	4.7	37
Rb/Sr	0.049	0.050	0.0036	0.0043	0.0036	0.0041	0.0092	0.0005
Sr/Zr	2.51	2.78	13.4	12.5	10.1	1.96	0.76	10.5
Sr/Ba	1.61	1.50	2.62	2.57	2.49	11.00	2.95	15.4
Zr/Nb	5.0	5.0	33.6	33.9	24.7	30.8	36.0	13.9
K/Rb	454	434	1,973	1,716	2,017	2,075	996	>3,000
La/Rb	1.2	1.1	3.9	3.4	4.8	5.5	>4.7	37
La/Nb	0.72	0.75	6.3	6.0	4.3	2.3	1.8	2.6
(La/Y) _N	7.3	7.2	21.7	22.8	23.9	1.92	0.87	20.5

* Guaymas Basin basalt (64/477A/2/2, 64–68 cm)¹⁴.† Basalt from mouth of Gulf of California (65/483/24/1, 96–100 cm)²⁴.‡ Magnesian andesite, Isla Cook, Chile³².§ Total ion as Fe₂O₃.|| Fe₂O₃/FeO = 0.2.¶ Measured Fe₂O₃ and FeO.

midden deposits dated at 5,000–6,000 yr BP (ref. 16) have a much greater vegetation cover than either the Jaraguay or San Borja lavas, suggesting an even younger age for the latter. Moreover, a lava underlying the Jaraguay flows discussed here has yielded a whole rock K–Ar age of 2.69 Myr (unpublished data). Petrographically, the lavas from San Quintin are distinct in having olivine as the dominant phenocryst phase with subordinate labradorite and rare clinopyroxene. Groundmasses contain olivine, magnetite, plagioclase microlites and augite. In contrast, most of the lavas from San Borja contain only sparse olivine crystals, and clinopyroxene is the major, and often only, phenocryst phase. Pilotaxitic groundmasses comprise plagioclase microlites, magnetite, clinopyroxene and an iron-rich mesostasis. The Jaraguay and a few San Borja flows are intermediate between the two in having dominant phenocryst olivine, subordinate clinopyroxene and trace plagioclase. Plagioclase microlites, clinopyroxene and an iron-rich mesostasis form the groundmass.

Representative analyses of the lavas from the three volcanic fields are presented in Table 1, together with samples drilled from the Guaymas Basin in the central part of the gulf of California¹⁴, and from the mouth of the gulf²⁴. The lavas from San Quintin are distinct from those of Jaraguay and San Borja, being nepheline-normative alkali basalts whose elemental variations reflect fractional crystallization of olivine + Cr spinel²². Zr/Nb (4.5–5.4), K/Rb (410–482) and La/Nb (0.66–0.75) ratios are lower than N-type mid-ocean ridge basalts (MORB) but comparable to other suites of alkali basalts²⁵. Chondrite-normalized La/Y ratios ((La/Y)_N) of 6.8–7.8 indicate slight light rare-earth elements (LREE) enrichment, but there is no marked

depletion of the heavy rare-earth elements (HREE). In contrast, rocks from Jaraguay are olivine or quartz normative basaltic andesites. The most striking geochemical feature is their high Sr contents (2,068–2,152 p.p.m.) which results in low Rb/Sr (0.0031–0.0043) and high Sr/Zr (12.5–13.5) ratios. Rb is also depleted with respect to K₂O, the K/Rb ratios varying from 1,716 to 2,314 (Fig. 2). Furthermore, all the large-ion lithophile elements (LILE) are enriched relative to the high field strength elements (HFSE) (for example, La/Nb = 5.2–6.7). Inter-HFSE ratios, however, have MORB-like affinities (for example, Zr/Nb = 29.3–37.0; Hf/Ta = 13.8–15.1). The lavas are strongly LREE enriched (Fig. 3) with (La/Yb)_N = 45.7–47.9 and (La/Y)_N = 17.5–22.8; moreover, Yb and Y show pronounced depletion with Yb_N = 3.5–3.6 and Y_N = 3.95–4.75. The high MgO contents of 6.2% at 55% SiO₂ classify the rocks as magnesian basaltic andesites; high Ni (150–157 p.p.m.) and Cr (199–221 p.p.m.) concentrations preclude much low-pressure fractionation.

The Holocene lavas from the San Borja volcanic field are also magnesian basaltic andesites with MgO ~5.0% at 55% SiO₂, these rocks, however, have undergone low-pressure fractionation of olivine + clinopyroxene + Cr spinel as indicated by their generally lower Ni and Cr contents (68–140 p.p.m. and 144–242 p.p.m. respectively). They show features of the Jaraguay lavas; high Sr (2,163–2,609 p.p.m.) and Ba (946–1161 p.p.m.) concentrations, high Sr/Zr (9.5–10.7), K/Rb (1,484–2,177) and La/Nb (4.2–5.9), and low Rb/Sr (0.0035–0.0046) ratios. (La/Y)_N ratios (18.4–23.9) are very fractionated, with Y_N = 4.3–6.7. Zr/Nb ratios (23.3–39.5) overlap those of Jaraguay but are often a little more enriched (Fig. 2).

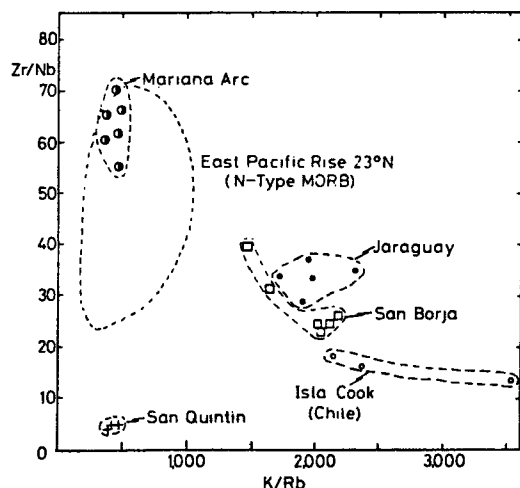


Fig. 2 Zr/Nb versus K/Rb for Holocene lavas from San Quintin, Jaraguay and San Borja volcanic fields of Baja California Norte and Isla Cook, southern Chile. MORB from EPR 23°N (ref. 24) and arc lavas from the Marianas⁴ are shown for comparison.

The depletion of Y and HREE in the San Borja and Jaraguay lavas attests to the importance of garnet as a residual phase during melting, assuming that the source was not already HREE depleted. MORB-normalized plots for the three lava types (Fig. 4) reveal that whereas San Quintin is similar to other within-plate alkali basalts^{5,25}, Jaraguay and San Borja have pronounced Nb depletions relative to LILE, a characteristic of subduction-related magmatism⁴⁻⁶. Interestingly, recent basalts recovered from the Guaymas Basin¹⁴ and from Isla Tortuga¹⁸ (a tholeiitic seamount), both in the central part of the Gulf of California, also portray enrichments of LILE compared with those from the gulf mouth where there is active spreading²⁰. It thus appears that this enrichment in alkalis and LREE is inherited from the subduction regime which was operative beneath north-west Mexico 12.5 Myr ago, and is superimposed on a MORB-type mantle.

High-magnesian andesites have been reported from other localities²⁶⁻³⁷. Those from the Bonin Islands, Mariana Forearc and Trench and Cape Vogel²⁶⁻³⁰ have been ascribed to the boninite series. They differ from the Jaraguay and San Borja lavas in having much higher MgO (>9% at 55% SiO₂) and much lower TiO₂ and Zr contents, higher K/Rb and lower Rb/Sr ratios and very depleted REE concentrations (<20 times chondrite and usually <10). Petrographically boninites are also distinct in having phenocrysts of orthopyroxene or clinopyroxene, while lacking olivine. The magnesian andesites from Baja California are thus not members of the boninite series. However, magnesian andesites similar to Jaraguay and San Borja have been described from several volcanic arcs³¹⁻³⁵. Those from the Setouchi volcanic belt in Japan are considered to be primary melts from hydrous mantle³⁴⁻³⁷. In the Aleutians, Kay concluded that the highly fractionated REE profiles (and, incidentally, the high Sr contents) of magnesian andesites (Fig. 3) were produced by eclogite melting. However, this cannot be the case in Baja as the enrichment of Sr and the depletion of Rb are too extreme. Magnesian andesites recently reported from Isla Cook in southernmost Chile^{32,33} bear a close resemblance to the Jaraguay and San Borja samples, particularly in their high Sr and very low Rb contents (Table 1, Fig. 3). They have depleted Sr and Nd isotope systematics (⁸⁷Sr/⁸⁶Sr = 0.70268–0.70280; ¹⁴³Nd/¹⁴⁴Nd = 0.513136), 'mantle' $\delta^{18}\text{O}$ (5.1–5.7) and Pb isotopes similar to other volcanic centres in the southern Andes^{33,38}; they are thought to have risen through the crust with little interaction. It is not possible, with the available data, to preclude crustal contamination during the formation of the Jaraguay and San Borja lavas. It is, however, unlikely. First, the low Rb content and concomitant very low Rb/Sr, Rb/Ba and Rb/La ratios imply that assimilation of sialic upper crust has been minimal. Second, the low ⁸⁷Sr/⁸⁶Sr ratio (0.7037) reported

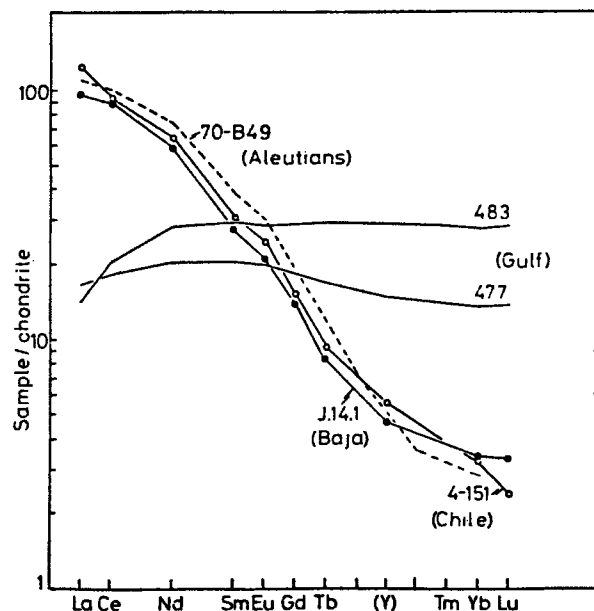


Fig. 3 Chondrite-normalized REE plot for the lavas from Jaraguay (J14.1) (unpublished data); the Gulf of California (Site 483: Gulf Mouth²⁴; Site 477: Guaymas Basin¹⁴; see Table 1); Isla Cook, Chile (4-151)³²; and western Aleutian Ridge (70-3 49)³¹. All analyses by instrumental neutron activation analysis.

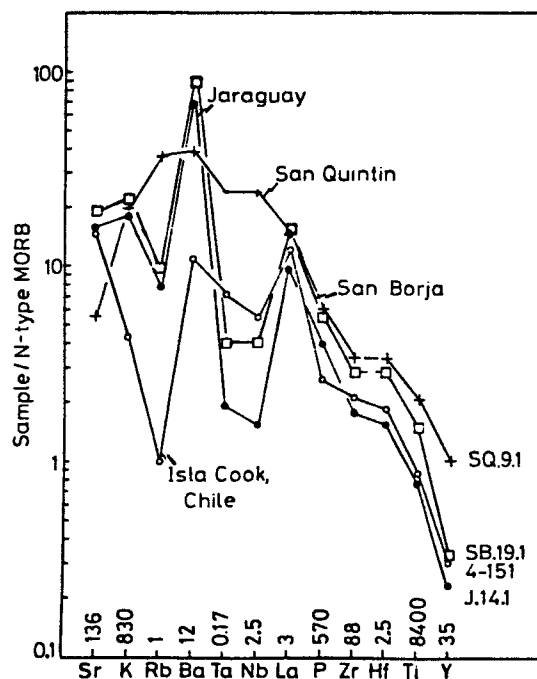


Fig. 4 MORB-normalized trace element plot for the Holocene lavas of Baja California Norte and Isla Cook, Chile³² (Table 1). MORB-normalizing values from ref. 43.

for the Jaraguay lavas⁹ precludes significant contamination either of the source by subducted sediment, or of the ascending magma. The isotope data for Jaraguay, Aleutians and southern Chile also rule out incorporation of pelagic, calcareous sediments in the magma sources to explain the very high Sr contents of these lavas.

Most interesting, however, is the apparent tectonic association of these magnesian andesites with the subduction of an actively spreading oceanic ridge. Isla Cook, for example, is sited on a transcurrent fault zone at the junction of the South American, Antarctic and Scotia plates where subduction has virtually ceased following the northward migration of the Chile Rise–Peru–Chile Trench triple junction in response to the oblique subduction of the Chile Rise. The magnesian andesites from the

Aleutians have an estimated middle Miocene age³¹, which overlaps the final stages of the effects of subduction of the Kula Ridge³⁹. The Setouchi high-magnesian andesites were erupted 12–14 Myr ago³⁷, a date which succeeds cessation of subduction of the Shikoku Basin crust at the Nankai Trough, during which process there was consumption of part of the marginal basin spreading centre⁴⁰. Finally, the Baja Holocene lavas were erupted 12.5 Myr after subduction of the EPR. Although Moore⁴¹ considers that the Baja California peninsula has moved north-west by ~240 km during the past 4 Myr, this movement has been along the Pacific margin and thus over regions of the mantle into which the ancestral EPR had already been subducted.

The association of these magnesian andesites (here termed Bajaite series rocks) with ridge subduction may result from increased thermal input into the mantle beneath the arc⁴² and/or from modification of the chemical input from the subducting slab compared with periods of 'normal' subduction of oceanic lithosphere. It is evident that volcanism associated with ridge subduction has a distinctive geochemistry, namely the generation of magnesian andesites with fractionated REE profiles, low Rb/Sr and high K/Rb, Sr/Zr, La/Rb and La/Nb ratios superimposed on a MORB-type HFSE signature. The mechanism for producing such compositions is unclear, but probably involves garnet as a residual phase during melting of a source variably enriched in Sr, and depleted in Rb, relative to most island arc magma sources. The transcendent fault-related setting of Baja California and Isla Cook may allow magmas to rise rapidly through the crust with insignificant contamination and only minor fractionation.

We thank J. Tarney, N. J. Norry and I. W. Luff for constructive criticism. Samples from Isla Cook were provided by M. Suarez. The study is supported financially by NERC grant GR3/4957.

Received 11 January; accepted 27 March 1985.

- Gill, J. B. *Orogenic Andesites and Plate Tectonics* (Springer, Berlin, 1981).
- Hawkesworth, C. J. in *Orogenic Andesites and Related Rocks* (ed. Thorpe, R. S.) 549–571 (Wiley, Chichester, 1982).
- Saunders, A. D., Tarney, J. & Weaver, S. D. *Earth planet. Sci. Lett.* **46**, 344–360 (1980).
- Hole, M. J., Saunders, A. D., Marriner, G. F. & Tarney, J. *J. geol. Soc. Lond.* **141**, 453–472 (1984).
- Pearce, J. A. in *Orogenic Andesites and Related Rocks* (ed. Thorpe, R. S.) 525–548 (Wiley, Chichester, 1982).
- Pearce, J. A. in *Continental Basalts and Mantle Xenoliths* (eds Hawkesworth, C. J. & Norry, M. J.) 230–249 (Shiva, Nantwich, 1983).
- Atwater, T. M. *Bull. geol. Soc. Am.* **81**, 3513–3536 (1970).
- Atwater, T. M. & Molnar, P. *Stanford Univ. Publ. geol. Sci.* **13**, 136–148 (1973).
- Dickinson, W. R. & Snyder, W. S. *J. geol.* **87**, 609–627 (1979).
- Mammerikx, J. & Klitgord, K. D. *J. geophys. Res.* **87**, 6751–6759 (1982).
- Sawlan, M. G. & Smith, J. G. *Pacif. Sect. S.E.P.M.* **39**, 237–251 (1984).
- Karig, D. E. & Jensen, W. *Earth planet. Sci. Lett.* **17**, 169–174 (1972).
- Larson, R. L. *Bull. geol. Soc. Am.* **83**, 3345–3360 (1972).
- Saunders, A. D., Fornari, D. J., Joron, J.-L., Tarney, J. & Treuil, M. *Init. Rep. DSDP* **64**, 595–642 (1982).
- Saunders, A. D., Fornari, D. J. & Morrison, M. A. *J. geol. Soc. Lond.* **139**, 335–346 (1982).
- Gastil, R. G., Phillips, R. P. & Allison, E. C. *Mem. geol. Soc. Am.* **140**, (1975).
- Gastil, R. G., Krummenacher, D. & Minch, J. *Bull. geol. Soc. Am.* **90**, 839–857 (1979).
- Batiza, R. *Bull. geol. Soc. Am.* **89**, 1309–1324 (1978).
- Mooser, F. & Rayes, L. A. *Mex. Univ. Nac. Inst. Geol. Bol.* **61**, 47–48 (1961).
- Sawlan, M. G. *Geol. Soc. Am. Abstr. Prog.* **14**, 231 (1982).
- Woodford, A. O. *Am. J. Sci.* **15**, 337–345 (1928).
- Stroh, J. M. thesis, Univ. Washington (1975).
- Basu, A. R. & Murthy, V. R. *Earth planet. Sci. Lett.* **35**, 239–246 (1977).
- Saunders, A. D. *Init. Rep. DSDP* **65**, 591–621 (1983).
- Chen, C.-Y. & Frey, F. A. *Nature* **302**, 785–789 (1983).
- Jenner, G. A. *Chem. Geol.* **30**, 307–332 (1981).
- Meijer, A., Anthony, E. & Reagan, M. *Init. Rep. DSDP* **60**, 709–729 (1981).
- Wood, D. A., Marsh, N. G., Tarney, J., Joron, J.-L., Fryer, P. & Treuil, M. *Init. Rep. DSDP* **60**, 611–645 (1981).
- Kuroda, N., Shiraki, K. & Urano, H. *Bull. Volcan.* **41**, 563–575 (1978).
- Hickey, R. L. & Frey, F. A. *Geochim. cosmochim. Acta* **46**, 2099–2115 (1982).
- Kay, R. W. *J. Volcan. geotherm. Res.* **4**, 117–132 (1978).
- Puig, A., Hervé, M., Suarez, M. & Saunders, A. D. *J. Volcan. geotherm. Res.* **20**, 149–163 (1984).
- Stern, C. R., Futa, K. & Muehlenbachs, K. in *Andean Magmatism Chemical and Isotopic Constraints* (eds Harmon, R. S. & Barreiro, B. A.) 31–46 (Shiva, Nantwich, 1984).
- Tatsumi, Y. & Ishizaka, K. *Earth planet. Sci. Lett.* **53**, 124–130 (1981).
- Tatsumi, Y. & Ishizaka, K. *Lithos* **15**, 161–172 (1982).
- Tatsumi, Y. *Earth planet. Sci. Lett.* **60**, 293–304 (1982).
- Tatsumi, Y. & Ishizaka, K. *Earth planet. Sci. Lett.* **60**, 293–304 (1982).
- Barreiro, B. A. in *Andean Magmatism Chemical and Isotopic Constraints* (eds Harmon, R. S. & Barreiro, B. A.) 21–30 (Shiva, Nantwich, 1984).
- DeLong, S. E. & Fox, P. J. in *Island Arcs, Deep Sea Trenches and Back-Arc Basins* (eds Talwani, M. & Pitman, W. C.) 221–228 (American Geophysical Union, 1977).
- Ingle, J. C. *Ir et al. Init. Rep. DSDP* **31**, 275–316 (1975).
- Moore, D. G. *Bull. geol. Soc. Am.* **84**, 1883–1906 (1973).
- DeLong, S. E., Schwarz, W. M. & Anderson, R. N. *Earth planet. Sci. Lett.* **44**, 239–246 (1979).
- Saunders, A. D. & Tarney, J. *Spec. Publ. geol. Soc. Lond.* No. 16, 59–76 (1984).

Cenozoic weathering covers in Buchan, Scotland and their significance

A. M. Hall

Fettes College, Carrington Road, Edinburgh EH4 1QX, UK

Isolated occurrences of chemically weathered rock or saprolite are common in many formerly glaciated areas around the North Atlantic^{1–6}. These pockets of saprolite are widely regarded as remnants of Cenozoic deep-weathering covers which existed before the first extensive Northern Hemisphere glaciations at ~2.4 Myr (ref. 7) and their apparent survival has long been used as evidence for the local ineffectiveness of glacial erosion^{2,4,8,9}. However, the scarcity of correlative deposits means that the age and significance of these fragmented saprolites are uncertain. I report here the results of a detailed survey which shows that in Buchan, Scotland, a combination of Cenozoic tectonic stability and extremely limited Pleistocene glacial erosion has allowed the preservation of deep-weathering covers of probable Miocene to early Pleistocene age on a scale not previously reported from any other formerly glaciated area.

Buchan is an area of subdued relief lying at the angle of north-east Scotland (Fig. 1). The geology consists largely of Precambrian to Dalradian metamorphic rocks intruded by Caledonian acid and basic igneous masses¹⁰. The only sedimentary rocks onshore are Devonian Old Red Sandstones but thick accumulations of Mesozoic and Palaeogene sediments lie offshore.

In Buchan, saprolites are continuous over wide areas¹¹ and maximum depths exceed 60 m in boreholes^{12,13} (Fig. 2). All major rock types are affected and alteration decreases with depth, indicating that weathering is a product of subaerial rather than hydrothermal alteration.

Within major lithologies, saprolite characteristics vary widely and reflect important differences in the degree of chemical alteration. Data on granulometry and clay mineralogy, backed by geochemistry, indicate that two main weathering types exist. Most widespread is the gruss weathering type in which a relatively low degree of alteration has caused disintegration into a granular sand. Gruss is characterized by an abundance of little-altered primary minerals, low fines contents, low soluble base losses and heterogeneous and immature clay mineral assemblages (Table 1). The second weathering type, clayey gruss, is virtually confined to central Buchan. Here advanced alteration has reduced parent rocks to a sticky, clayey silty sand. Detrital primary mineralogy is dominated by quartz, fines contents are relatively high, and clay mineral assemblages consist of mature kaolinite-illite suites. At several sites, clayey grusses are rubefied due to the presence of haematite and goethite. This distinction between grusses and clayey grusses is comparable with others made for granite weathering covers in Europe south of the limits of Pleistocene glaciation^{5,14,15}.

The relative ages of the two weathering types can be established using evidence from correlative deposits, saprolite mineralogy and geomorphology. All deep weathering in north-east Scotland must predate the last glaciation as materials reworked from subjacent and adjacent saprolites form a major component of Devensian tills and congeliturbates^{11,16,17}. Evidence that many gruss weathering profiles are of middle Pleistocene age or older is provided by: (1) the presence of weathered biotite and feldspar and inherited clays throughout the middle to late Pleistocene sequence of glacial, glaciifluvial and periglacial deposits at Kirkhill^{18,19}; and (2) the presence of hydromorphic iron and manganese minerals in weathering profiles now occupying free-draining sites²⁰. These minerals indicate that weathering had already extended close to and probably below the water table before a major change in groundwater drainage status during the Pleistocene in response to regional drainage incision.

Fig. 1 Location and relief.

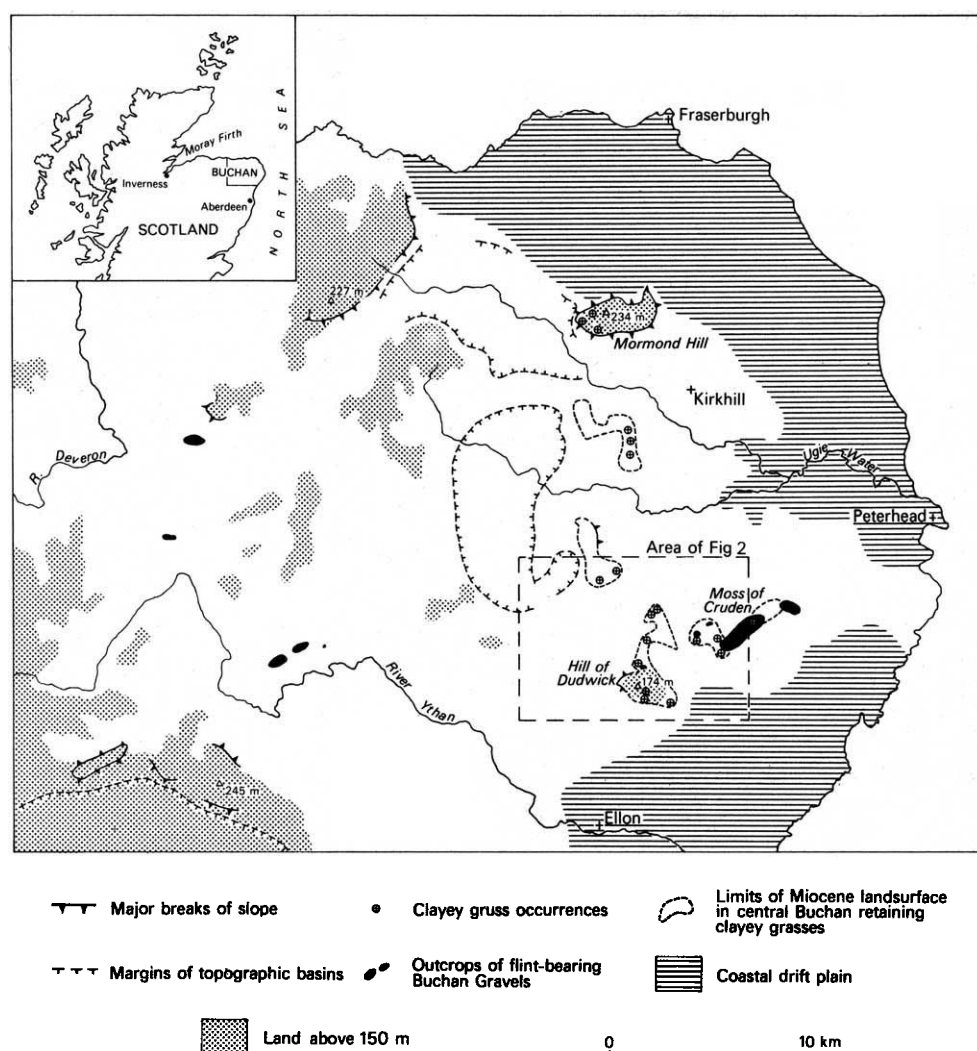


Table 1 Summary characteristics of weathering types

Rock type	Granite						Basic igneous			Quartz schist					Pelite	
Site and grid reference	Hill of Longhaven NK 084423	Cairnlea NJ 901537	Blackrigg NJ 912600	Mill Maud NJ 566067	Backhill NK 026407	Pittodrie NJ 693245	Silverford NJ 432253	Fedderate NJ 888505	Cruden borehole NK 028402	Moss-side NJ 949568	Howford NJ 954547	Mormond Quarry NJ 950568	Sunnyside NJ 983372	Howe of Dens NJ 975805	Kinnuck NJ 991341	Windy Hills borehole NJ 793393
Weathering type	G	G	G	G	G	CG	G	G	CG	G	G	CG	CG	CG	G	CG
% < 2 μ m	0.2	1.9	1.7	2.1	3.6	17.0	0.2	1.8	43.5	2.1	8.2	12.0	15.4	13.7	7.1	22.2
% < 63 μ m	2.5	14.4	12.7	17.5	26.1	45.0	6.7	11.3	75.5	9.1	41.3	32.0	43.0	49.8	20.1	94.7
Feldspar/quartz ratio	1.2	0.9	1.2	1.8	1.1	0.1	ND	ND	ND	0.5	0.7	0.1	0.1	0.2	ND	ND
% MgO	0.36	0.36	0.44	0.34	0.53	0.11	6.60	4.52	0.49	0.51	1.23	0.24	0.15	0.14	ND	0.18
% CaO	1.34	0.08	0.12	0.14	0.37	0.07	8.44	4.66	0.49	0.20	0.23	0.07	0.07	0.07	ND	0.10
% Na ₂ O	3.36	1.13	1.58	2.14	1.81	0.14	1.98	1.59	0.05	1.22	0.18	0.15	0.10	0.05	ND	0.55
% K ₂ O	6.03	5.3	4.65	5.63	3.56	3.00	0.56	0.55	0.45	2.71	3.84	1.24	0.63	2.80	ND	0.72
% Soluble base losses	5.9	53.4	58.6	48.6	ND	95.4	19.0	31.1	ND	ND	ND	ND	ND	ND	ND	78.9
Rubefaction	—	—	—	—	—	X	—	—	—	—	—	X	X	X	—	—
Etching of quartz	—	—	—	—	—	—	—	—	—	—	—	X	X	—	—	—
Clay mineralogy	I K H C	H	I	K I C	K H V S	K I Hm	I V	Gi K I	K I	Gi I K	I K H V	K I	K I Hm	K I Go	M C I	K I

C, Chlorite; Gi, gibbsite; Go, goethite; H, halloysite; I, illite; M, mixed-layers; S, smectite; V, vermiculite; G, gruss; CG, clayey gruss; ND, not determined.

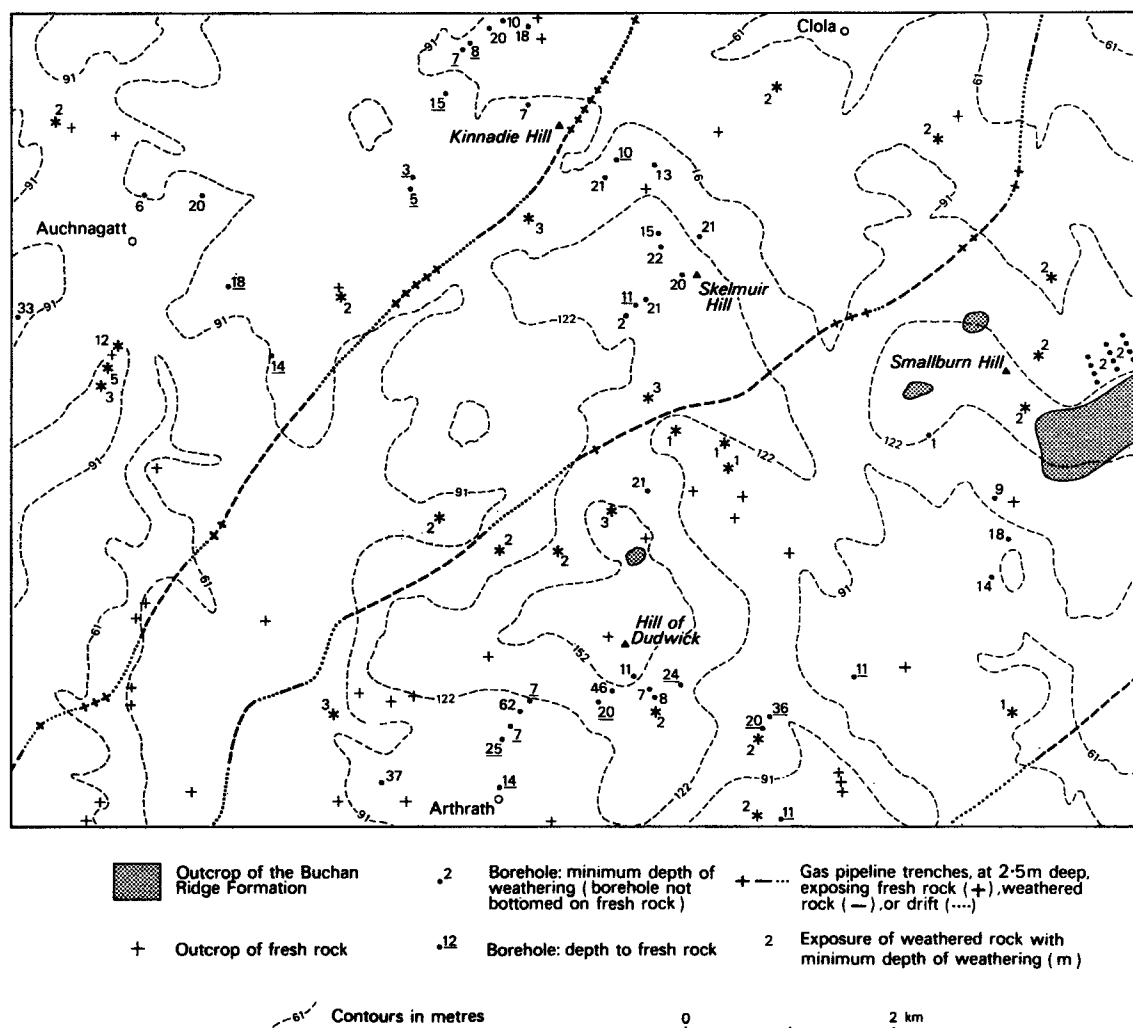


Fig. 2 Deep weathering patterns in central Buchan.

Suggestions that deep weathering is wholly of interglacial age²¹ may also be rejected in view of the disparity between rates of rock alteration during the present interglacial², the relatively short durations (~10–30 kyr) of previous interglacials²² and observed depths of weathering. Formation before the glacial Pleistocene is indicated.

The dominance of kandite and illite clay mineral assemblages in acid igneous and metamorphic grusses²⁰ and hydrous mica-amorphous iron oxide assemblages in basic igneous grusses¹⁶ indicates weathering under humid temperate environments. However, the intensity of alteration shown by the clayey grusses requires warmer, probably subtropical climates^{14,15}. Marine palaeotemperatures²³ and terrestrial palynology²⁴ indicate that subtropical climates last prevailed in the Scottish area during the middle Miocene (10 Myr). Major climatic cooling after this time^{23,24} is reflected in the mineralogy of North Sea well samples, with appearance of chlorite and amphibole and increases in illite and feldspar contents^{25,26}, and in a transformation in styles of rock weathering throughout western Europe with the first development in the Tertiary of sandy saprolites retaining significant amounts of detrital feldspar and biotite^{5,14}. I conclude that the clayey gruss weathering type predates the late Miocene and that the gruss weathering type developed continuously throughout the Pliocene and early Pleistocene.

The clayey grusses are virtually confined to a north-south belt of high ground in central Buchan (Fig. 1) which represents the remnants of a Miocene or older landscape. Around Hill of Dudwick, the clayey grusses are juxtaposed with and underlie outcrops of the Buchan Ridge Formation. These intensely-

weathered gravel deposits comprise clasts of rolled Cretaceous flint and Dalradian quartzite with a matrix of quartz sand and kaolinitic fines²⁷, and are widely regarded as Tertiary in age^{28–30}. At Moss of Cruden, the gravels are >25 m thick and include clasts of little-travelled nodular flint. The preservation of such large volumes of flint suggests that Buchan has remained close to base-level since marine transgression in the late Cretaceous. This reflects the position of Buchan at the stable eastern extremity of the Scottish Highlands, an area repeatedly tilted westwards towards the rapidly subsiding North Sea basin in the Tertiary³¹.

Tectonic stability has been a major factor in the extraordinary development and state of preservation of Cenozoic weathering covers. However, the survival of friable saprolites also demonstrates the singular ineffectiveness of Pleistocene glacial erosion in Buchan, where a succession of cold-based ice sheets have done little to modify Neogene landscapes¹⁹.

Pockets of gruss-type saprolites occur widely in other glaciated areas bordering the North Atlantic^{1–6}. Analogies with Buchan suggest that during the Neogene, lowland crystalline terrains were extensively weathered under temperate environments, with depths of weathering exceeding 30 m on interfluvies. Where saprolites have survived glaciation and can be shown not to be the products of hypogene alteration³² or exceptionally deep alteration down fault zones^{12,20}, it can be proposed that local Pleistocene glacial erosion has amounted to at most the removal of a few tens of metres of regolith and exposure of the basal surface of weathering. Such estimates can be reconciled with regional evidence of deep Pleistocene erosion deduced from the

volumes of offshore sediments^{33,34} by recognizing the selectivity of glacial erosion. Enclaves of little-modified Neogene relief have survived in parts of eastern Canada^{4,35}, eastern Greenland³⁶, Fennoscandia^{5,37-39} and the British Isles^{2,9,40} where local factors, including remoteness from ice accumulation centres, low gradients and diffluent flow patterns, have combined to maintain basal ice below its pressure melting point and so prevent significant glacial erosion³⁶. Finally, the thicknesses of saprolite reported here would suggest that material reworked from preglacial Cenozoic weathering profiles forms a major, but largely overlooked component of glacial tills^{37,41}.

This work was carried out during tenure of a Research Scholarship from the University of St Andrews. A grant towards the cost of illustrations was provided by the Carnegie Trust for the Universities of Scotland. Borehole information was provided by the Department of Geology, University of Aberdeen, RioFinEx Ltd and the British Geological Survey, Edinburgh.

Received 24 January; accepted 3 April 1985.

1. Goldthwait, J. W. & Kruger, F. C. *Bull. geol. Soc. Am.* **49**, 1183-1198 (1938).
2. Godard, A. *Recherches géomorphologiques en Écosse du Nord-ouest* (Masson et Cie, Paris, 1965).
3. Walsh, P. T., Boulter, M. C., Ijtaba, M. & Urbani, D. M. *J. geol. Soc. Lond.* **128**, 519-559 (1972).
4. Gauthier, R. C. *Geol. Surv. Can. Curr. Res. Pap.* 80-1B, 277-282 (1980).
5. Lidmar-Bergström, K. *Sver. geol. Unders. Ser. C* No. 785 (1982).
6. Roaldset, E., Pettersen, E., Longva, O. & Mangerud, J. *Norsk geol. Tidsskr.* **62**, 169-178 (1982).
7. Shackleton, N. J. *et al. Nature* **307**, 620-623 (1984).
8. Feininger, T. *U.S. geol. Surv. Prof. Pap.* 750C, 65-81 (1971).

9. Battiau-Queney, Y. *Rev. Géom. Dyn.* **30**, 63-73 (1981).
10. Johnstone, G. S. *British Regional Geology: The Grampian Highlands* (HMSO, Edinburgh, 1966).
11. FitzPatrick, E. A. *J. Soil. Sci.* **14**, 33-42 (1963).
12. Edmond, J. M. & Graham, J. D. *Q.J. Engng Geol.* **10**, 281-301 (1977).
13. Leslie, A. G. *Scott. J. Geol.* **20**, 215-235 (1984).
14. Bakker, J. P. & Levett, Th. W. M. *Publ. Serv. Geol. Lux.* **14**, 27-75 (1964).
15. Millot, G. *The Geology of Clays* (Chapman & Hall, London, 1970).
16. Basham, I. R. *Clay Miner.* **10**, 189-202 (1974).
17. Wilson, M. J. & Tait, J. M. *Clay Miner.* **12**, 59-66 (1977).
18. Connell, E. R., Edwards, K. J. & Hall, A. M. *Nature* **297**, 570-572 (1982).
19. Hall, A. M. (ed.) *Buchan Field Guide* (Quaternary Research Association, Cambridge, 1984).
20. Hall, A. M. thesis, Univ. St Andrews (1983).
21. Phemister, T. C. & Simpson, S. *Nature* **164**, 318-319 (1949).
22. Bowen, D. Q. *Quaternary Geology* (Pergamon, Oxford, 1978).
23. Buchardt, B. *Nature* **275**, 121-123 (1978).
24. Mudie, P. J. & Helgason, J. *Nature* **303**, 689-692 (1983).
25. Karlsson, W., Vollset, J., Bjorlykke, K. & Jorgensen, P. *Proc. 6th int. Clay Conf.* **27**, 281-289 (1979).
26. Berstad, S. & Dypvik, H. *J. Petrol. Geol.* **5**, 77-88 (1982).
27. Kesel, R. H. & Gemmell, A. M. D. *Scott. J. Geol.* **17**, 185-203 (1981).
28. Koppi, A. J. & FitzPatrick, E. A. *J. Soil. Sci.* **31**, 525-532 (1980).
29. McMillan, A. A. & Merritt, J. *Rep. Inst. geol. Sci. No. 80/1* (1980).
30. Hall, A. M. *Scott. J. Geol.* **18**, 336-338 (1982).
31. Ziegler, P. A. in *Petroleum Geology of the Continental Shelf of North-West Europe* (eds Illing, L. V. & Hobson, G. D.) 3-39 (Heyden, Rhine, 1981).
32. Samuelsson, L. *Sver. geol. Unders. Arsh.* **67**(9), 16 (1973).
33. Matthews, W. H. *Am. J. Sci.* **275**, 818-824 (1975).
34. Laine, E. D. *Quat. Res.* **14**, 188-192 (1980).
35. McKeague, J. A., Grant, D. R., Kodama, H., Beke, G. J. & Wang, C. *Can. J. Earth Sci.* **20**, 37-48 (1983).
36. Sugden, D. E. *Inst. Br. Geogr. Spec. Publ.* **7**, 177-196 (1974).
37. Kaitenen, V. *Fennia* **99**(5), 85 (1969).
38. Kiselev, I. I. *Izv. vses. geogr. Obshch.* **107**, 133-148 (1975).
39. Soderman, G., Kejonen, A. & Kujansuu, R. *Fennia* **161**, 91-144 (1983).
40. Mitchell, G. F. *J. Earth Sci. (Dubl.)* **3**, 13-33 (1980).
41. Rosenquist, I. Th. *Geol. För. Stockh. Förh.* **97**, 284-286 (1975).

Isotope composition of sulphates generated by bacterial and abiological oxidation

Robert O. van Everdingen

National Hydrology Research Institute, 4616 Valiant Drive, Calgary, Alberta, Canada T3A 0X9

H. Roy Krouse

Physics Department, The University of Calgary, Calgary, Alberta, Canada T2N 1N4

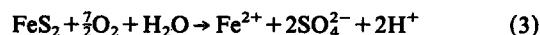
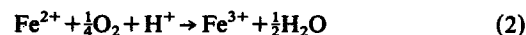
Oxygen isotope compositions of secondary sulphates provide a potentially powerful tool for elucidating mechanisms of sulphur oxidation. Such studies should reveal the consequences of various anthropogenic sulphur contributions to the environment, including gaseous emissions, aqueous solutions and elemental sulphur dust. At mining sites, metal sulphides may be oxidized *in situ*, giving rise to acid mine waters. The environment can receive these compounds from natural sources also, for example volcanoes, sulphurous springs and sour-gas seeps. There are many fundamental questions concerning the specific oxidation reactions and associated kinetics. In principle, isotope data from appropriate experiments can be used to address such questions. In the case of acid mine drainage, the approach of Taylor *et al.*¹ is noteworthy. Here we present a critical examination of their results and suggest a novel method of determining the kinetic isotope effects identified with the incorporation of O₂ and H₂O-oxygen by SO₄²⁻.

Oxygen isotope data may provide information about redox reactions in the sulphur cycle when sulphur isotope data fail to do so. For example, SO₄²⁻ in subsurface water may be partially reduced to HS⁻ in deeper anaerobic environments, and the reduced product may be subsequently re-oxidized near the surface. The sulphur isotope composition may be conserved if processes such as metallic sulphide deposition or H₂S escape are absent; in this case there would be no sulphur-isotope evidence of the reduction-re-oxidation sequence. However, during re-oxidation, oxygen incorporated in the product SO₄²⁻ can be derived from both O₂ and H₂O molecules; this generally

produces δ¹⁸O values different from those of the original SO₄²⁻. The above scenario has been documented for sulphates in several springs in western Canada²⁻⁴.

Taylor *et al.*¹ presented a scheme whereby a plot (their Fig. 2) of δ¹⁸O values of SO₄²⁻ versus those of associated H₂O can be used to evaluate the relative participation of two specific oxidation reactions. In our laboratory, ¹⁸O/¹⁶O analyses have been done on sulphate produced by oxidation of biogenic H₂S (near sulphurous springs)^{2,3} and of metal sulphides (in acidic springs)². It would seem that the approach of Taylor *et al.*¹ could be applied to these data. However, when this approach was attempted, it became apparent that their Fig. 2 required refinement.

Taylor *et al.*¹ cited three reactions identified with the oxidation of FeS₂:



In reaction (1) all sulphate oxygen (O_s) is derived from water molecules, whereas in reaction (3) 87.5% of the sulphate oxygen is derived from molecular oxygen and 12.5% from water molecules. Sulphates produced by the two reactions have different δ¹⁸O_s values because of the generally large difference between the δ¹⁸O values of O₂ and H₂O (δ¹⁸O_a and δ¹⁸O_w, respectively).

Even if reactions (1) and (3) apply, there are two reasons why the isotope composition of oxygen in the sulphate does not necessarily reflect the stoichiometric proportions. First, a reactant atom need not end up in the product; it may instead enter a solvent molecule or another ion, depending on the reaction mechanisms and isotope exchange reactions. Second, reaction (2), the rate-limiting step for reaction (1), generates H₂O with a δ¹⁸O value similar to that of O₂, which may affect the isotope balance of reaction (1). In the absence of data bearing on these possibilities, the stoichiometry of reactions (1) and (3) will be assumed to apply to the isotope composition, as done by Taylor *et al.*¹

In their Fig. 2, Taylor *et al.*¹ plotted δ¹⁸O values for SO₄²⁻ in acid mine waters (δ¹⁸O_s) against those for associated H₂O

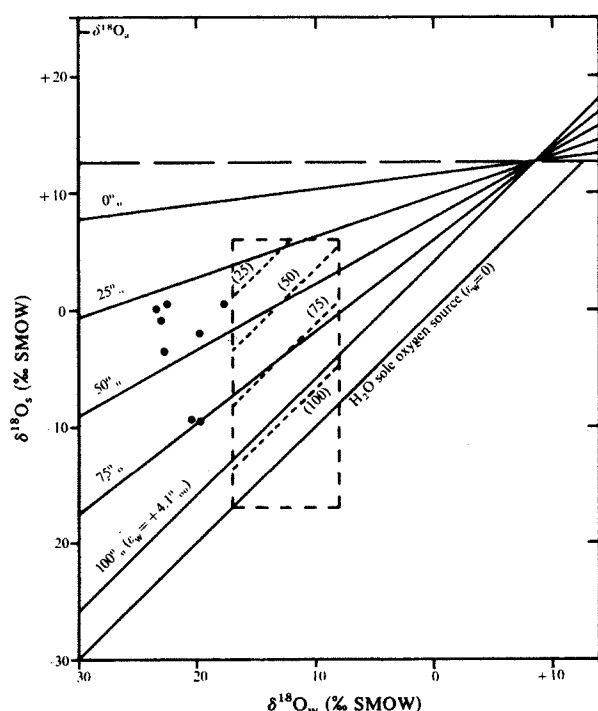


Fig. 1 Expected $\delta^{18}\text{O}_s$ values for secondary sulphates produced by oxidation of pyrite plotted against the $\delta^{18}\text{O}_w$ value of associated water, for relative contributions by reaction (1) ranging from 0 to 100%. Broken lines are from Taylor *et al.*¹. Points represent data for acidic springs in western Canada.

($\delta^{18}\text{O}_w$). They drew a line representing zero difference between $\delta^{18}\text{O}_s$ and $\delta^{18}\text{O}_w$, labelled 'H₂O sole O₂ source', and a parallel line, offset by about +3.5‰ and labelled 100%, representing sulphates produced exclusively by reaction (1). The offset reflects the estimated kinetic isotope effect for the reaction. Three further lines, corresponding to 75, 50 and 25% contribution by reaction (1) were drawn parallel to the first two lines.

When we plotted $\delta^{18}\text{O}_s$ values for dissolved sulphates from acidic spring waters in western Canada, associated with lower $\delta^{18}\text{O}_w$ values (−23‰ standard mean ocean water, SMOW) than the data of Taylor *et al.*¹, an unrealistic contribution from reaction (3) was indicated. On further reflection, it was realized that the lines labelled 75, 50 and 25% should not be parallel. This can be readily appreciated by assuming a hypothetical reaction whereby all oxygen in the product sulphate is derived from O₂. In that case, the plot would be a horizontal line, as $\delta^{18}\text{O}_s$ would be independent of $\delta^{18}\text{O}_w$. Whereas the line for 100% contribution of sulphate oxygen by reaction (1) should be parallel to the line for 'H₂O sole oxygen source' and have a slope of unity (Fig. 1), decreasing contributions by reaction (1) would produce lines with progressively decreasing slopes.

The above considerations are incorporated in the isotope balance equation

$$\delta^{18}\text{O}_s = Y(\delta^{18}\text{O}_w + \epsilon_w) + (1 - Y)[0.875(\delta^{18}\text{O}_a + \epsilon_a) + 0.125(\delta^{18}\text{O}_w + \epsilon_w)] \quad (4)$$

in which Y is the fraction of SO_4^{2-} produced by reaction (1); ϵ_w and ϵ_a represent shifts in $\delta^{18}\text{O}_w$ and $\delta^{18}\text{O}_a$ values due to kinetic isotope effects during incorporation of O_w and O_a into sulphate. Using $\epsilon_w = +4.1\%$ (ref. 1), $\epsilon_a = -11.2\%$ (ref. 1) and $\delta^{18}\text{O}_a = +23.8\%$ (ref. 5), equation (4) can be rewritten as

$$\delta^{18}\text{O}_s = \delta^{18}\text{O}_w(0.875Y + 0.125) + (11.5375 - 7.4375Y) \quad (5)$$

Lines for 75, 50, 25 and 0% contribution by reaction (1), calculated using equation (5), are plotted in Fig. 1; they converge and intersect at the point where $\delta^{18}\text{O}_s = +12.6\%$ and $\delta^{18}\text{O}_w = +8.5\%$. This convergence means that at low latitudes, where

$\delta^{18}\text{O}_w$ values are closer to SMOW, it is more difficult to differentiate contributions from individual reactions than at higher latitudes, where $\delta^{18}\text{O}_w$ values may be lower than −30‰.

The 75, 50 and 25% lines from Fig. 2 of Taylor *et al.*¹ have been added to our Fig. 1, together with $\delta^{18}\text{O}$ data from acidic springs in western Canada. It is readily seen that extrapolation of the parallel lines of Taylor *et al.*¹ would indicate very different proportions for the relative contributions by reactions (1) and (3) to these sulphates.

Figure 1 also suggests that oxidation of FeS₂ in contact with water having a $\delta^{18}\text{O}_w$ value of +8.5‰ should produce sulphate having a $\delta^{18}\text{O}_s$ value equal to +12.6‰, independent of the relative contributions by reactions (1) and (3). It should be possible to verify this point experimentally; this in turn would provide a check on the values for ϵ_a and ϵ_w , as well as on the validity of the assumption that the stoichiometry of reaction (3) applies, or whether the overall approach is too simplified.

Received 29 November 1984; accepted 28 March 1985.

1. Taylor, B. E., Wheeler, M. C. & Nordstrom, D. K. *Nature* **308**, 538–541 (1984).
2. van Everdingen, R. O., Shakur, M. A. & Krouse, H. R. *Can. J. Earth Sci.* **19**, 1395–1407 (1982).
3. Shakur, M. A. thesis, Univ. Calgary, Alberta (1982).
4. van Everdingen, R. O., Shakur, M. A. & Krouse, H. R. 4th Int. Symp. Water–Rock Interaction, 526–529 (Okayama University Press 1983).
5. Horibe, Y., Shigehara, K. & Takakura, Y. *J. geophys. Res.* **78**, 2625–2629 (1973).

Fluorescent bands in massive corals result from terrestrial fulvic acid inputs to nearshore zone

Kevin Boto & Peter Isdale

Australian Institute of Marine Science, PMB No. 3, Townsville M.C., Queensland 4810, Australia

The discovery of yellow-green fluorescent bands in the skeletons of massive corals from inshore waters of the Great Barrier Reef, and the temporal and spatial correlations of these bands with adjacent river flow¹, prompted an investigation into the substances responsible for the phenomenon. We have determined that all regions of the coral skeleton contain fulvic acids^{2,3} which impart a blue background fluorescence. The yellow-green fluorescent bands found only in inshore coral appear to result from enhanced concentrations of low-relative molecular mass (<10,000) fulvic acids from adjacent river input. These compounds are ubiquitous, and the yellow-green banding phenomenon can be expected to occur in massive corals located near to significant river systems. In the widely distributed genus *Porites*⁴, the great age of the largest colonies¹ means that extremely long, highly resolvable records of coastal runoff exist throughout tropical nearshore seas.

Fluorescence spectra of coral core sections were obtained using a modified version of the apparatus described previously¹. The new version employed a 150-W mercury arc excitation beam, passed through a precision 360-nm bandpass filter and chopped at 285 Hz. A quartz optical fibre focused this beam into a 3-mm spot on the coral surface. Another optical fibre passed the fluorescence emitted from the coral surface into the monochromator and detector system of a Varian AA6 atomic absorption spectrophotometer. This instrument was tuned to the chopper frequency to eliminate interference from ambient light.

The fluorescence spectra for the banded and background regions of the coral (Fig. 1) are very broad, with maxima at 440–460 nm. These are indicative of humic and fulvic acids, mixtures of complex polyphenolic compounds with a wide range of relative molecular mass ($M_r < 5 \times 10^2 - 5 \times 10^4$). Fulvic acids are generally in the lower M_r range ($< 5 \times 10^2 - 3 \times 10^4$) and are defined as the fraction of the humic material soluble at pH 1 (ref. 2). The ubiquitous blue fluorescence found in marine and terrestrial waters³ is thought to result from the presence of these

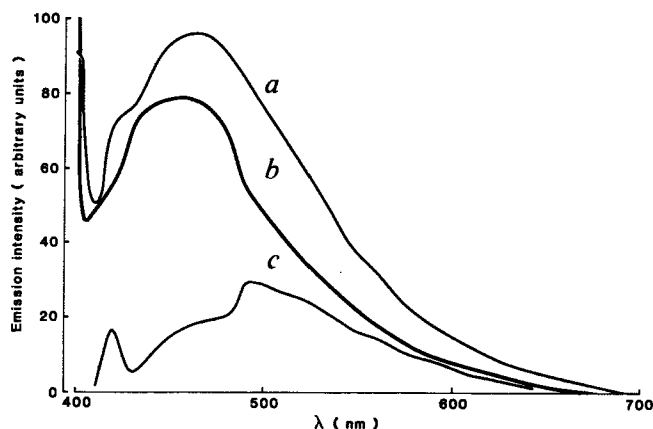


Fig. 1 Emission spectra of the: a, 'yellow-green band'; and b, background regions of an inshore *Porites* spp. coral skeleton. The 'yellow-green band' corresponds to the large January 1974 flow recorded in the nearby Burdekin River¹. c, Direct subtraction of b from a.

humic compounds. In particular, the high fluorescence efficiency and refractory nature of terrestrial humics has enabled their use as conservative tracers of fluvial input to coastal waters^{3,6}. The molecular structures and the nature of the fluorescing chromophore(s) of these compounds are poorly understood^{2,5}, although their general fluorescent characteristics and strong metal-complexing nature have been well documented^{7,8}.

The following simple experiments were performed to test the hypothesis that the fluorescence in the inshore-growing corals was a consequence of humic inclusions in the crystal structure of the skeleton. All solutions were adjusted to pH 6.5 before obtaining fluorescence spectra using an Aminco-Bowman spectrofluorometer.

(1) Dissolution of coral samples in dilute HCl gave solutions which showed broad fluorescence excitation and emission spectra typical of humic compounds⁵. Excitation (light absorption) maxima occurred in the 320–340-nm region with distinct shoulders at 250–270 nm. Emission maxima occurred at 410–440 nm. The greater wavelength of emission maxima for the intact corals (Fig. 1) is compatible with the greater degree of crosslinking and structural rigidity conferred on the humic molecules when contained in a solid matrix².

(2) Solutions of similar fluorescence characteristics to the above were obtained by hot alkaline (0.1 M NaOH) extraction of the crushed coral samples, denoting the existence in the skeleton of organic acid functional groups.

(3) Addition of excess Fe^{3+} to solutions obtained by acid dissolution or alkaline extraction of the coral shifted the emission maximum (+25 nm), with 30% reduction in intensity. These observations are consistent with the formation of a strong humic material- Fe^{3+} complex, and the quenching of fluorescence of these compounds by Fe^{3+} (ref. 3).

(4) Extracts of coral samples, acidified to pH 1, were passed through Amberlite XAD-2 resin which has a strong affinity for adsorption of humic compounds in the acid form⁹. It was found that 90% of the fluorescent components were adsorbed onto the resin and could be eluted subsequently with NH_4OH solution pH 11.

(5) All of the fluorescent material in the coral was apparently soluble at pH 1, as evidenced by unchanged total fluorescence after 0.2- μm filtration. Following the usual definition² we shall refer to this material as fulvic acid.

Comparison of the emission spectra (Fig. 1) for the background and 'yellow-green band' areas of the coral shows that their maximum differences occur in the 480–530-nm region. These differences, although small, combined with the human eye response around 540–560 nm, result in the yellow-green appearance of the bands and their remarkable visible contrast

Table 1 Apparent M_r distribution (% in each size class) of the dissolved organic carbon in acid extracts of the yellow-green fluorescent band and background regions of a coral skeleton¹

Apparent M_r	Yellow-green				Background			
	Distribution* (%)		Fluorescence Ex Em		Distribution* (%)		Fluorescence Ex Em	
<1,000	32.4	32.1	320	415	25.5	21.3	330	415
10^3 – 10^4	30.6	27.9	325	430	13.2	18.1	320	415
> 10^4	37.0	40.0	325	415	61.3	60.6	332	420

Extract solutions were adjusted to pH 6.5 and passed through a 0.2- μm filter before ultrafiltration using Amincon UM-2 (M_r 1,000) and UM-10 (10,000) membranes¹⁰. The fluorescence characteristics of the resulting fractions are expressed as Ex (wavelength of maximum excitation) and Em (wavelength of maximum emission (nm)).

* Results of duplicate analyses.

with the blue background fluorescence. We consider that the fulvic acid composition in the two distinct regions of the skeleton is sufficiently varied to impart these small but visually prominent differences in their emission characteristics. We investigated the M_r distribution of the compounds present in acid extracts of the two different regions of the skeleton. We employed Amincon ultrafilters with nominal M_r cutoffs of 10^3 and 10^5 daltons, using a similar procedure to that described by Cole *et al.*¹⁰. The dissolved organic carbon (DOC) content of initial, filtrate, and concentrate solutions was determined using a Graetzl DOC analyser¹¹. Table 1 shows that the yellow-green fluorescing region of the coral contained significantly greater ($P < 0.05$) proportions of fulvic acids in the < 10^3 and 10^3 – 10^5 daltons size classes than did the blue-fluorescing background region of the coral.

The synchrony of deposition of the yellow-green fluorescent bands with adjacent fluvial inputs, and the lack of such bands in mid- and outer-shelf corals¹ allows us to postulate that the enhanced fulvic acid levels in the yellow-green bands are of terrestrial origin. These terrestrial fulvics probably differ considerably in structure and chemical characteristics from the fulvic compounds generally present in sea water¹².

We were able to induce yellow-green fluorescence in the fast-growing branching coral *Acropora formosa* by incubating the colonies in local sea water to which fulvic acid extracted from a local alluvial soil had been added at either 10 or 15 mg l^{-1} . This demonstrated conclusively that the inclusion of these ambient compounds can impart the unusual fluorescence characteristics to coral skeletons.

The global distribution of massive species of *Porites* combined with the ubiquity of fulvic compounds and their ingress to nearshore environments via fluvial pathways, suggests that the fluorescent banding in inshore corals will be a general phenomenon. In accordance with previous arguments¹, it implies that centuries of rainfall/runoff analogues are available in corals from a global range of locations.

We thank our colleagues from the Australian Institute of Marine Science for assistance, especially D. Barnes, W. Bandaranayake, A. Nott, E. Daniel and M. Rennie. This is contribution no. 273 from the Australian Institute of Marine Science.

Received 18 January; accepted 2 April 1985.

- Isdale, P. *Nature* **310**, 578–579 (1984).
- Schnitzer, M. in *Soil Organic Matter* (eds Schnitzer, M. & Khan, S. U.) 1–64 (Elsevier, Amsterdam, 1978).
- Willey, J. C. *Mar. Chem.* **15**, 19–45 (1984).
- Veron, J. E. N. & Pichon, M. *Aust. Inst. mar. Sci. Monogr. Ser. V* (1982).
- Larson, R. A. & Rockwell, A. L. *Arch. Hydrobiol.* **89**, 416–425 (1980).
- Willey, J. D. & Atkinson, L. P. *Estuar. Coast. Shelf Sci.* **14**, 49–59 (1982).
- Stewart, A. J. & Wetzel, R. G. *Limnol. Oceanogr.* **25**, 559–564 (1980).
- Ghassemi, M. & Christman, R. F. *Limnol. Oceanogr.* **13**, 583–597 (1968).
- Stuermer, D. H. & Harvey, G. R. *Deep Sea Res.* **24**, 303–309 (1977).
- Cole, J. J., McDowell, W. H. & Likens, G. E. *Oikos* **42**, 1–9 (1984).
- Mueller, H. & Bandaranayake, W. *Mar. Chem.* **12**, 59–68 (1983).
- Hedges, J. I. & Parker, P. L. *Geochim. cosmochim. Acta* **40**, 1019–1029 (1976).

Distinctive mammal-like reptile from Mexico and its bearing on the phylogeny of the Tritylodontidae

James M. Clark & James A. Hopson

Department of Anatomy, University of Chicago, 1025 East 57th St, Chicago, Illinois 60637, USA

The skull of an unusual new mammal-like reptile of the family Tritylodontidae, recently discovered in the Jurassic of northeastern Mexico, is described here as *Bocatherium mexicanum*. Tritylodontids are the last survivors of the great radiation of mammal-like reptiles that also gave rise to mammals at or near the Triassic–Jurassic boundary (~200 Myr ago). These superficially rodent-like herbivores coexisted with early carnivorous mammals for some 50 Myr before becoming extinct in the late Jurassic. Although eight genera are recognized, their interrelationships have remained unclear because it has been impossible to determine the direction of morphological change within this morphologically stereotyped group. Unusual specializations of the Mexican skull now permit character polarities within the family to be determined and a cladogram to be constructed. *Bocatherium* shows a close relationship to *Bienotheroides*¹ of China and *Stereognathus*^{2,3} of Great Britain. It is the first mammal-like reptile and probably the oldest terrestrial vertebrate from Mexico.

Controversy remains as to the closest relatives of tritylodontids; they are either the terminal taxon of the gomphodont cynodonts, a group showing herbivorous adaptations similar to those of tritylodontids⁴, or very close relatives of the earliest true mammals, with which they share many similarities⁵. The eight genera described so far are: *Tritylodon*² and *Tritylodontoides*⁶ from southern Africa; *Bienotherium*⁷, *Lufengia*⁸ and *Bienotheroides*¹ from China; *Stereognathus*^{2,3} from Great Britain; *Oligokyphus* from western Europe⁹ and Arizona¹⁰; and *Kayentatherium*¹¹ from Arizona. *Stereognathus* is mid-Jurassic (Bathonian)³ and *Bienotheroides* is mid- or late Jurassic^{1,12} whereas the remaining taxa are considered to be of early Jurassic age¹³. Fragmentary postcranial remains from the Los Colorados Formation of Argentina, probably late Triassic (Norian), seem to represent the oldest-known tritylodontid¹⁴. We have been prompted to reassess the intrafamilial relationships of tritylodontids by the discovery of a nearly complete skull with lower jaws from northeastern Mexico (collected by J.M.C. in November 1982). The specimen pertains to a new genus and species which is described below.

Division Synapsida (Therapsida)¹⁵
Family Tritylodontidae

Table 1 Measurements of *B. mexicanum* type (mm)

Basal skull length Right cheek teeth	51.1	
	Length	Width
M1/	4.2	5.2
M2/	4.1	5.1
M3/	4.0	5.0
M4/	4.1	5.0
M/1	5.2	3.8
M/2	5.2	3.9
M/3	5.2	4.0

Genus *Bocatherium*, gen. nov.

Diagnosis. *Bocatherium* is unique among tritylodontids in having a large boss on the lateral surface of the dentary anterior to the masseteric fossa. It differs from other tritylodontids with the exception of *Bienotheroides* and *Stereognathus* in that: (1) its maxilla is reduced to a simple, nearly cylindrical bone lacking a medially directed horizontal lamina contributing to the secondary palate, a dorsally directed facial process in contact with the nasal, and a laterally directed lamina underlying the jugal on the anterior root of the zygomatic arch; (2) its upper post-canine teeth have only two cusps in each of three longitudinal rows (other tritylodontids have at least three cusps in the central row). *Bocatherium* is distinguished from *Bienotheroides* in having a slender, rather than extremely deep, zygomatic arch. It is distinguished from *Stereognathus* (known only from isolated teeth and a left maxilla with three teeth) in having upper post-canine teeth with an oval rather than quadrangular outline in crown view.

Bocatherium mexicanum, spec. nov.

Diagnosis. As for genus.

Type. IGM 3492, Instituto de Geología, Universidad Nacional Autónoma de México, México City; a skull with lower jaws, lacking the right zygomatic arch, the roof and much of the sides of the braincase.

Horizon and locality. Approximately the middle of the 400-m-thick exposed section of the La Boca Formation in Huizachal Canyon near the towns of Huizachal, Anacahuíta and La Joya, 15 km south-west of Ciudad Victoria, Tamaulipas State, México. The La Boca Formation consists of terrigenous sediments which lie between Permian and late Jurassic (Oxfordian) marine sediments^{16–18}. The occurrence of a tritylodontid suggests a Jurassic age for that part of the formation from which the fossil was collected. However, late Triassic plant fossils are reported from discontinuous beds assigned to the La Boca Formation at a locality 12 km to the north¹⁶. The conflicting ages probably indicate that the La Boca Formation contains strata of both

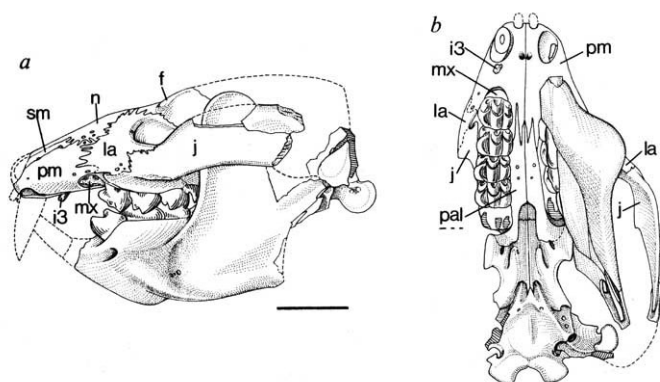


Fig. 1 *Bocatherium mexicanum*. Holotype skull (IGM 3492) in lateral (a) and ventral (b) views. Right mandible removed in b. Key: f, frontal; i, incisor; j, jugal; la, lacrimal; mx, maxilla; n, nasal; pal, palatine; pm, premaxilla; sm, septomaxilla. Scale bar, 1 cm.

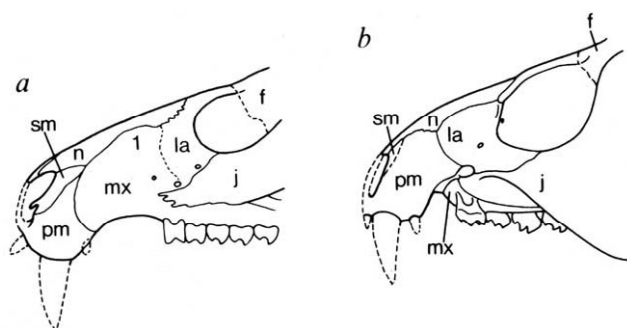


Fig. 2 Lateral views of tritylodontid facial regions representing the primitive pattern of facial bones exemplified by *Bienotherium* (a) and the derived pattern exemplified by *Bienotheroides*¹ (b). Not to scale. Key: 1, the facial process of the maxilla; other abbreviations are as in Fig. 1.

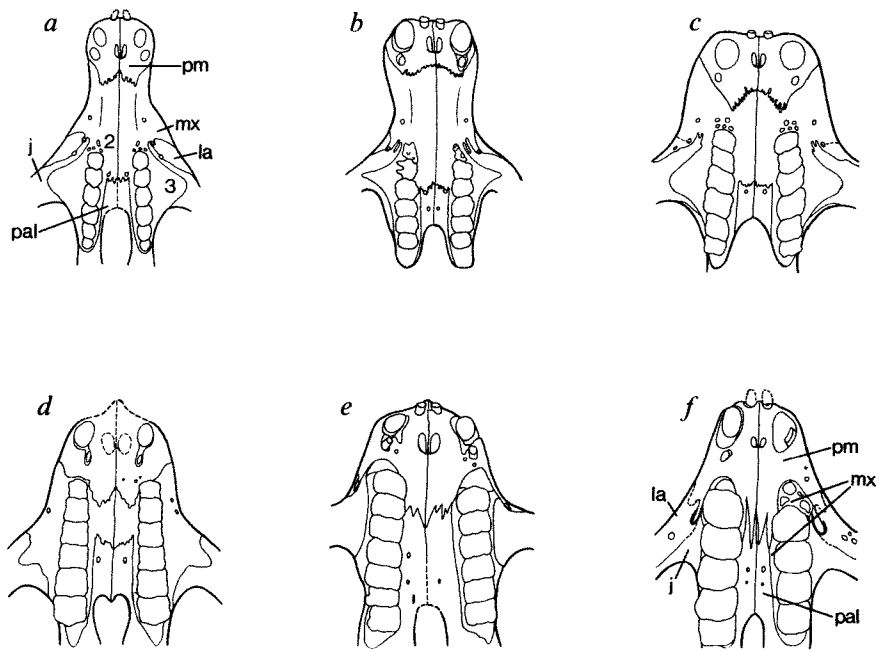


Fig. 3 Ventral views of tritylodontid skulls. *a*, *Oligokyphus*⁹; *b*, *Tritylodon*; *c*, *Bienotherium*; *d*, *Kayentatherium* (H.-D. Sues, unpublished); *e*, *Bienotheroides*¹; *f*, *Bocatherium*. Key: 2, palatine process of maxilla; 3, zygomatic process of maxilla; other abbreviations as in Fig. 1. Cusps on teeth omitted. Not to scale.

Triassic and Jurassic age, as has already been suggested by a palaeomagnetic study of rocks in this area, which identified three diachronous units within the La Boca Formation¹⁹. The Huizachal Canyon beds have the youngest pole position of the three, but a pole position has not been determined for rocks of the plant locality.

Bocatherium mexicanum (Fig. 1) is one of the smallest tritylodontids (Table 1), only slightly larger than *Lufengia minor*⁸, but the only known skull may represent a sub-adult individual. Immaturity is suggested by the unusual depth and robustness of the dentary and by the relatively small number and large size of the cheek teeth, features characterizing juvenile but not adult *Bienotherium yunnanense* (J.A.H., unpublished). However, unlike immature individuals of other tritylodontids (see, for example, Fig. 3*d*), the teeth are of equal size rather than the posterior teeth becoming increasingly larger (Table 1). In overall morphology, *Bocatherium* resembles other tritylodontids, differing most notably in the bulbous swelling on the dentary anterior to the masseteric fossa. Basicranial and braincase structure, as preserved, are typically tritylodontid^{9,20}.

In the pattern of bones on the side of the face and on the palate, *Bocatherium* is strikingly different from all other tritylodontids (Figs 2*a*, 3), indeed, from all other cynodonts (cladistically including mammals), except *Bienotheroides*¹ and *Stereognathus* (J.A.H., unpublished). In these three genera, the maxilla is reduced to a nearly cylindrical element holding the roots of the teeth; thus, it lacks the usual thin laminae extending upward onto the face, medially onto the secondary palate and laterally below the jugal. *Bocatherium* and *Bienotheroides* (Figs 2*b*, 3), and by inference *Stereognathus*, show the following specialized features. The premaxilla is greatly enlarged so that it is in contact with the nasal and the anteriorly expanded lachrymal on the side of the face and the anteriorly expanded palatine on the secondary palate. The jugal is expanded antero-posteriorly on the lateral surface of the maxilla to form a long flat vertical contact. This contact facet is seen on the type specimen of *Stereognathus ooliticus*, usually described as a right maxilla^{2,3} but actually a left (J.A.H., unpublished). Opening onto the face at the junction of the jugal, premaxilla and lachrymal is a large infraorbital foramen, the medial wall of which is formed by a prominent groove on the maxilla (visible in *S. ooliticus*).

The crown pattern of the upper postcanine teeth of these three genera differs from that of all other tritylodontids in having two

cusps in each of three longitudinal rows, with very small accessory cusps at the anterior end of each row. The crowns are oval in occlusal outline in *Bocatherium* and *Bienotheroides*, quadrangular in *Stereognathus*.

Knowledge of the extremely derived condition of the facial and palatal bones in these three genera permits the remaining genera known from adequate skull material to be arranged in a morphocline (Fig. 3). (*Tritylodontoides*⁶ is too poorly preserved to be included.) The most derived early Jurassic taxa are *Kayentatherium* (H.-D. Sues, unpublished) and *Lufengia* (J.A.H., unpublished), in which the premaxilla and palatine closely approach each other on the palate so that the maxillary component is greatly reduced. The premaxilla is in contact with the nasal in *Lufengia* juveniles (known adult skulls are damaged in this region) but not in *Kayentatherium* and other early Jurassic genera (Fig. 2*a*). The palatine of *Bienotherium* shows the primitive unexpanded condition, but the premaxilla extends well back

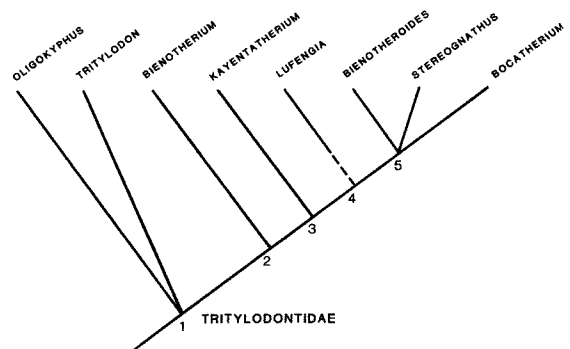


Fig. 4 Cladogram showing proposed relationships of tritylodontid genera. Synapomorphies for each node are as follows: 1, Second upper and first lower incisors enlarged; first upper incisor very small; canines absent; cheek teeth with divided roots; postcanines with three (upper) or two (lower) longitudinal rows of crescentic cusps; postorbital bar and prefrontal and postorbital bones absent. 2, Premaxilla extends further posteriorly on palate; snout relatively short with no post-incisive constriction. 3, Premaxilla extends posteriorly between anterior postcanines, and palatine extends anteriorly beyond level of palatine foramina so that palatine process of maxilla is reduced. 4, Premaxilla in contact with nasal, and facial process of maxilla reduced in area (in *Lufengia* juveniles; adult condition uncertain). 5, Facial, palatine and zygomatic processes of maxilla absent; upper cheek teeth with two principal cusps in each longitudinal row.

on the palate (J.A.H., unpublished). All of the above genera are similar in possessing the presumably derived condition of a short snout lacking a constriction behind the incisor region. *Tritylodon* (J.A.H., unpublished) and *Oligokyphus*⁹ have a short premaxilla and a long narrow snout with a mid-length constriction; these features are found both in early mammals²¹ and in gomphodont cynodonts (J.A.H., unpublished) and thus are considered primitive for tritylodontids. The interrelationships of tritylodontid genera (Fig. 4) suggest that a small number of cusps in the upper postcanine teeth is a derived condition.

The close relationship of *Bocatherium* to *Bienotheroides* and *Stereognathus* suggests that the Mexican form may also be post-early Jurassic in age; it is certainly younger than late Triassic. A Jurassic age for the upper part of the La Boca Formation is consistent with the stratigraphical¹⁸ and palaeomagnetic¹⁹ evidence. We believe *Bocatherium* to be the oldest terrestrial vertebrate known from Mexico. If the upper part of the La Boca Formation proves to be middle Jurassic in age, *Bocatherium* will be the first terrestrial vertebrate of this age in North America^{22,23}.

We thank R. C. Belcher, O. Carranza, R. Cifelli, J. C. Guerrero, A. C. Milner, H.-D. Sues, A.-L. Sun and M. O. Woodburne for their aid. This research was supported by grants from the NSF (to J.A.H.) and the Hinds Fund of the University of Chicago

and the Theodore Roosevelt Fund (to J.M.C.). Illustrations were prepared by C. Vanderslice.

Received 27 December 1984; accepted 20 March 1985.

1. Sun, A.-L. *Scientia Sinica* B1984, 257-268 (1984).
2. Simpson, G. G. *A Catalogue of the Mesozoic Mammals in the Geological Department of the British Museum* (British Museum (Natural History), London, 1928).
3. Waldman, M. & Savage, R. J. G. *J. geol. Soc.* 128, 119-125 (1972).
4. Crompton, A. W. *Bull. Br. Mus. nat. Hist.* A21, 27-71 (1972).
5. Kemp, T. S. *Zool. J. Linn. Soc.* 77, 353-384 (1983).
6. Fourie, S. *Navors. nas. Mus., Bloemfontein* 2, 7-19 (1962).
7. Young, C. C. *Proc. zool. Soc. Lond.* 117, 537-597 (1947).
8. Chow, M. C. & Hu, C. C. *Vertebr. palasiat.* 3, 9-12 (1959).
9. Kühne, W. G. *The Liassic Therapsid Oligokyphus* (British Museum (Natural History), London, 1956).
10. Jenkins, F. A. Jr, Crompton, A. W. & Downs, W. R. *Science* 222, 1233-1235 (1983).
11. Kermack, D. M. *Zool. J. Linn. Soc.* 76, 1-17 (1982).
12. Sigogneau-Russell, D. & Sun, A.-L. *Geobios* 14, 275-279 (1981).
13. Olsen, P. E. & Galton, P. M. *Science* 197, 983-986 (1977).
14. Bonaparte, J. F. *Op. Iilloana* 22, 1-183 (1971).
15. Wiley, E. O. *Syst. Zool.* 28, 308-337 (1979).
16. Mixon, R. B., Murray, G. E. & Diaz, T. *Bull. Am. Ass. Petrol. Geol.* 43, 757-771 (1959).
17. Carrillo Bravo, J. *Boln Asoc. Mex. Geol. Petrol.* 13, 1-98 (1961).
18. Imlay, R. W. *U.S. geol. Surv. prof. Pap.* 1062, 1-134 (1980).
19. Gosc, W. A., Belcher, R. C. & Scott, G. R. *Geology* 10, 50-54 (1982).
20. Hopson, J. A. *Postilla* 87, 1-30 (1964).
21. Kermack, K. A., Mussett, F. & Rigney, H. W. *Zool. J. Linn. Soc.* 71, 1-158 (1981).
22. Russell, D. A. *Sylogus* 53, 1-35 (1984).
23. Clemens, W. A., Lillegraven, J. A., Lindsay, E. H. & Simpson, G. G. in *Mesozoic Mammals* (eds Lillegraven, J. A., Kielan-Jaworowska, Z. & Clemens, W. A.) 7-58 (University of California, Berkeley, 1979).

Neo-darwinian evolution implies punctuated equilibria

C. M. Newman*, J. E. Cohen†§ & C. Kipnis‡

* Department of Mathematics, University of Arizona, Tucson, Arizona 85721, USA

† Laboratory of Populations, Rockefeller University, New York, New York 10021-6399, USA

‡ Centre de Mathématiques appliquées, Ecole Polytechnique, Palaiseau 91128, France

The two central elements of neo-darwinian evolution^{1,2} are small random variations and natural selection. In Wright's view, these lead to random drift of mean population characters in a fixed, multiply peaked 'adaptive landscape', with long periods spent near fitness peaks^{3,4}. Using recent theoretical results⁵, we show here that transitions between peaks are rapid and unidirectional even though (indeed, because) random variations are small and transitions initially require movement against selection. Thus, punctuated equilibrium^{6,7}, the palaeontological pattern^{8,9} of rapid transitions between morphological equilibria, is a natural manifestation of the standard wrightian evolutionary theory and requires no special developmental, genetic or ecological mechanisms¹⁰⁻¹³.

The simplest neo-darwinian model for the evolution of the population mean $\bar{x}$ of a genetically determined character expresses the change $d\bar{x}$ in time interval dt as the sum of a natural selection term and a random variation term:

$$d\bar{x}(t) = F'(\bar{x}(t)) + \alpha dW(t) \quad (1)$$

$F(\bar{x})$, a one-dimensional genotypic³ or phenotypic¹⁴ adaptive landscape, describes the mean population fitness. If, for example, the slope F' is positive (a rising landscape) at $\bar{x}(t)$, natural selection pushes $\bar{x}$ towards larger values. The parameter α gives the magnitude of random variations relative to that of natural selection. The random process $W(t)$, a standard brownian movement with zero mean drift, represents, for example, genetic drift and short-term environmental fluctuations showing no obvious trend. We assume, for simplicity, that α is independent of $\bar{x}$.

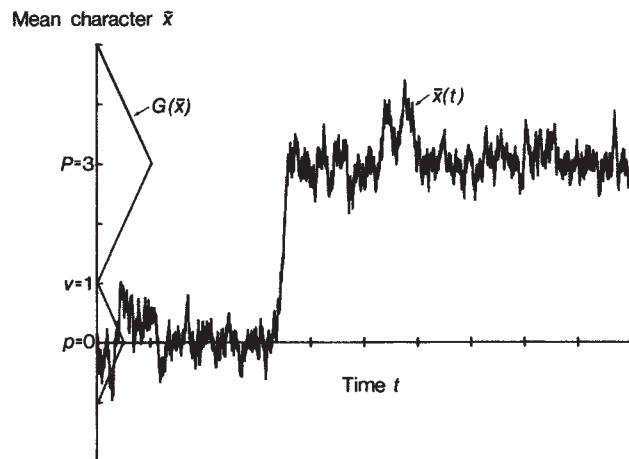


Fig. 1 Punctuated equilibria in a numerical solution of the discrete time equation, $\bar{x}(t+1) - \bar{x}(t) = G'(\bar{x}(t)) + sw_t$. The landscape G has a lower peak at $\bar{x}=0$, a valley at $\bar{x}=1$ and a higher peak at $\bar{x}=3$; its slope, G' , takes only the values $+0.01$ or -0.01 . w_t represents the sign of random variables, s their magnitude. For each t , w_t is independently $+1$ or -1 with equal probability; $s=0.07$. The jagged line plots $\bar{x}(t)$ for 5,000 time units. As theory predicts, during the transition between peaks, $\bar{x}$ moves from 0 to 1 as if the direction of natural selection were reversed, that is, at the same speed that $\bar{x}$ moves by natural selection from 1 to 3.

Suppose that the landscape is doubly peaked: as $\bar{x}$ increases, F rises to a peak at $\bar{x}=p$, declines to a valley at $\bar{x}=v$, rises to a second higher peak at $\bar{x}=P$, then declines. Let the initial value $\bar{x}(0)$ be smaller than v . To mimic the fossil record, suppose $\bar{x}(t)$ is observed, not continuously, but at some finite (perhaps large) number N of epochs $0 < t_1 < \dots < t_N$, perhaps chosen by another random process modelling stratigraphic sampling. There is a natural timescale τ_α associated with transitions from the smaller to the larger peak which increases dramatically as α decreases. We suppose that the intervals between observations are, for small values of α , proportional to τ_α .

Under these assumptions⁵, for small α , the observed $\bar{x}(t_i)$ values behave quite differently from $\bar{x}(t)$ itself, which is intrinsically continuous. As α approaches 0, the observed values become fixed exactly at p for some random (exponentially distributed) time, then jump (apparently) instantaneously to P .

§ To whom correspondence should be addressed at the Rockefeller University.

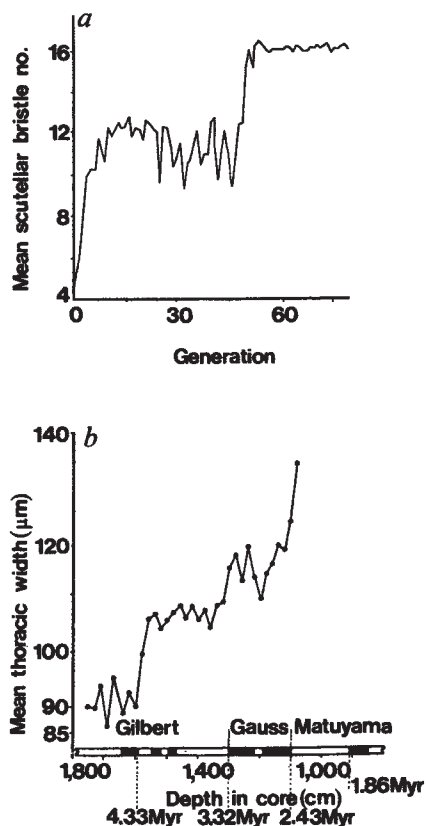


Fig. 2 *a*, Response to selection for scutellar bristle number in female *Drosophila* (presumably *melanogaster*; species not given¹⁶) according to MacBean *et al.*¹⁶, as simplified by Parsons¹⁸. *b*, Increase in thoracic width of the Antarctic radiolarian *Pseudocubus vema* during ~2.5 Myr according to Kellogg¹⁷, as simplified by Parsons¹⁸.

and remain there (apparently) indefinitely. Analogous results hold if F has more than two peaks⁵ or if $\bar{x}$ is multidimensional¹⁵.

To understand this result, note that a movement of $\bar{x}$ from p to v goes against the influence of natural selection and, for small α , is very improbable. Improbable events eventually occur (with timescale τ_α) and then by the most likely of the improbable alternatives. Here it is most likely⁵ that small random variations combine to push $\bar{x}$ from p to v as if the direction of natural selection were reversed, hence rapidly and unidirectionally. Once at v , the move to P is by natural selection. The rapidity of the second part of the transition is well recognized⁴, whereas that of the first part seems largely unknown.

Figure 1 is a computer simulation illustrating the theoretical results about equation (1). For comparison, Fig. 2 shows examples^{16,17} (from ref. 18) of observed punctuated equilibria with two different timescales.

In a changing landscape, punctuation can be triggered by selection, rather than drift, either by a rapid (for example, environmental) change¹⁹ or after a gradual change eliminates the currently occupied peak¹³. The introduction of a rare selectively superior mutant into a population from which it was previously absent can lead to punctuation^{20,21} corresponding to transition from an unstable equilibrium in a fixed landscape. Our analysis is the first to show that transitions between stable equilibria, triggered by drift in an unchanging adaptive landscape, should have fossil records indistinguishable from punctuations triggered by selection.

Gradual evolution is only possible in an unchanging landscape either if α is not small (violating neo-darwinian assumptions) or if sampling intervals are as short as the timescale of natural selection (unusual in fossil records). Even assuming many peaks, the suggestion^{20,22} that stepwise changes will appear to be gradual on a coarse timescale seems to be inconsistent

with Wright's model: an increased timescale of observation implies⁵ that more local peaks are skipped over apparently instantaneously and stasis is observed only at higher fitness levels.

Gradual evolution is possible in a changing landscape if peaks shift gradually and unidirectionally. Our analysis suggests a resulting palaeontological pattern of 'punctuated shifting equilibria' as the gradual evolution of a population following a moving peak is punctuated by rapid transitions between peaks. It is unclear whether such changing landscapes should occur commonly.

Models similar to that described in equation (1) are applicable directly to populations that are panmictic or have much migration and hence are geographically homogeneous. Because the timescale τ_α is a rapidly increasing function of population size, equation (1) predicts stasis and punctuation for a small to moderate population but only stasis for a large population.

The model in equation (1) may be applied to an isolated subpopulation with a smaller τ_α than the larger parent population. Allopatric speciation will occur if the subpopulation, after transition to a new peak, cannot interbreed with the parent population. Without requiring special 'genetic revolutions',²³ our analysis predicts that punctuated equilibria will be associated with the branching speciation of isolated subpopulations.

Generalizations of equation (1) incorporating spatial inhomogeneity in a local mean character correspond to Wright's notion of local populations in his shifting balance model^{3,4}, in which he suggested that speciation could occur without physical barriers to migration and gene flow²³. Mathematically non-rigorous^{24,25} and rigorous²⁶ results suggest that punctuation occurs on the timescale τ needed to form by random drift (against natural selection) a 'critical droplet' (that is²³, a geographical region where transition to a higher peak of fitness occurs) which is of sufficient size to resist swamping by gene flow; such a critical droplet spreads rapidly by selection. The timescale τ is a slowly decreasing function of population size²⁷ (for fixed spatial density) because a critical droplet can form in each local subrange^{4,23}. Unlike τ_α , this τ need not exceed geological timescales even for very large populations.

Our research was partially supported by USNSF grants MCS80-19384 (C.M.N.) and BSR84-07461 (J.E.C.), a Guggenheim Fellowship (C.M.N.), a MacArthur Fellowship (J.E.C.) and the hospitality of Mr and Mrs W. T. Golden (J.E.C.). We thank J. F. Crow, W. G. Faris, J. Felsenstein, L. R. Ginzburg, S. J. Gould, J. Maynard Smith, E. Mayr, W. M. Schaffer and L. S. Schulman for comments on earlier drafts.

Received 26 November 1984; accepted 12 March 1985.

- Mayr, E. *Animal Species and Evolution* (Harvard University Press, 1963).
- Lewontin, R. C. *The Genetic Basis of Evolutionary Change* (Columbia University Press, New York, 1974).
- Wright, S. *Proc. 6th int. Congr. Genet.* Vol. 1, 356-366 (1932).
- Wright, S. *Evolution and the Genetics of Populations* Vol. 3, 443-473; Vol. 4, 460-476 (University of Chicago Press, 1977, 1978).
- Kipnis, C. & Newman, C. M. *SIAM J. appl. Math.* (in the press).
- Eldredge, N. & Gould, S. J. in *Models in Paleobiology* (ed. Schopf, T. J. M.) 82-115 (Freeman, San Francisco, 1972).
- Gould, S. J. & Eldredge, N. *Paleobiology* 3, 115-151 (1977).
- Simpson, G. G. *Tempo and Mode in Evolution* (Columbia University Press, New York, 1944).
- Simpson, G. G. *The Major Features of Evolution* (Columbia University Press, New York, 1953).
- Hoffman, A. *Evol. Biol.* 15, 411-436 (1982).
- Maynard Smith, J. A. *Rev. Genet.* 17, 11-25 (1983).
- Geiger, G. *Math. Biosci.* 67, 59-79 (1983).
- Kirkpatrick, M. *Am. Nat.* 119, 833-848 (1982).
- Lande, R. *Evolution* 30, 314-334 (1976).
- Galves, A., Olivieri, E. & Vares, M. E. *Ann. Probabil.* (submitted).
- MacBean, I. T., McKenzie, J. A. & Parsons, P. A. *Theor. appl. Genet.* 41, 227-235 (1971).
- Kellogg, D. E. *Paleobiology* 1, 359-370 (1975).
- Parsons, P. A. *The Evolutionary Biology of Colonizing Species* (Cambridge University Press, 1983).
- Wright, S. in *Genetics, Paleontology and Evolution* (eds Jepsen, G. L., Mayr, E. & Simpson, G. G.) 365-389 (Princeton University Press, 1949).
- Ginzburg, L. *Paleobiology* 7, 426-429 (1981).
- Petry, D. *Paleobiology* 8, 56-66 (1982).
- Bookstein, F. L., Gingerich, P. D. & Kluge, A. G. *Paleobiology* 4, 120-134 (1978).
- Barton, N. H. & Charlesworth, B. A. *Rev. Ecol. Systematics* 15, 133-164 (1984).
- Fisher, M. E. *Physics* 3, 255-283 (1967).
- Langer, J. S. *Ann. Phys.* 41, 108-157 (1967).
- Faris, W. G. & Jona-Lasinio, G. *J. Phys. A* 15, 3025-3055 (1982).
- Newman, C. M. & Schulman, L. S. *J. statist. Phys.* 23, 131-148 (1980).

Interpolation in stereoscopic matching

G. J. Mitchison* & S. P. McKee†

*MRC Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH and Kenneth Craik Laboratory, Department of Physiology, University of Cambridge, Cambridge CB2 3EG, UK

†Smith-Kettlewell Institute for Visual Sciences, 2232 Webster Street, San Francisco, California, USA

Anyone who has stared at a repeating wallpaper pattern, or a periodic pattern of tiles, has probably experienced the phenomenon of a false stereoscopic depth percept. This arises because of a mismatching in the two eyes of repeating elements in the pattern. The phenomenon is less likely to occur if an edge of the textured region is in view; the edge seems to fix the registration of elements. We describe here a stereogram which exemplifies this principle; it has a central, periodic region bounded on either side by edges with pre-assigned disparities. We find that the perceived depth of the central region is controlled by the edges. In certain conditions (when the period is spatially large), the edges simply impose one of the expected discrete matchings. In other conditions, however, we observe a striking phenomenon: interpolation in depth occurs between the edges, violating any possible feature-by-feature matching.

Our stereogram can be broadly described as a regular grid of points with horizontal disparities introduced at the two vertical edges (Fig. 1A). A single row of the grid is shown in Fig. 1B, where the convention is that the vertical alignment of the dots seen by the left and right eyes corresponds to the zero disparity; that is, if each dot were matched to that vertically above or below, they would be seen in the fixation plane. The row consists of a set of equally spaced dots, typically 6 arc min apart, but with the left-most dot moved inwards a certain fraction, S_L , of the inter-dot spacing in the left eye, and the right-most dot moved inwards by a fraction S_R of the inter-dot spacing in the right eye.

The stereoscopic percept depends on how points in the two eyes are matched. When the stereogram is inspected carefully, using a large inter-dot spacing (for example, Fig. 2 viewed close up), one of two possible percepts is seen, depending on the size of S_L and S_R . When S_L and S_R are small, the percept corresponds to the matching shown in Fig. 1C(1), and consists of a plane with two flaps tilted forwards at the edges [Fig. 1D(1)]. When S_L and S_R are large, the underlying matching is that of Fig.

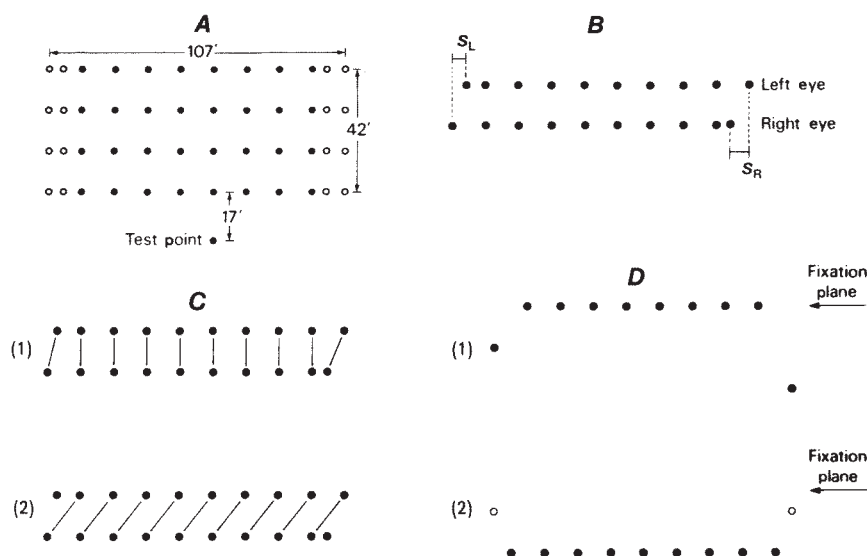
1C(2), in which each point is paired obliquely; the percept here is of a grid lying in front of the fixation plane, with two unpaired points at the edges [Fig. 1D(2)]. Thus, changing S_L and S_R causes a change of matching throughout the central regular lattice. Such propagation of matching from a limited area throughout a periodic region has been shown to occur in stereograms studied by Julesz and Chang¹, of which ours can be regarded as a highly specialized case.

To explore the depth percepts more systematically, we added a test point beneath the grid (Fig. 1A) and asked subjects, in a forced-choice paradigm, to judge whether the test point lay in front of, or behind, the grid. The test point was placed at a fixed vertical position below the grid, and could be given any desired horizontal location, so that depths of particular regions of the grid could be explored. The test point was randomly given one of seven equally spaced disparities, chosen so that at one extreme the grid lay considerably in front of the test point, and at the other, behind it. By fitting a cumulative gaussian curve to the percentage of trials on which the grid was reported as lying in front of the test point, we could locate the apparent mean position of the region of the grid above the test point, and also determine the standard deviation in its perceived position². Observers were first shown a fixation pattern which disappeared during the subsequent 160-ms presentation of the grid and test point.

When the (horizontal) inter-dot spacing in the central region was $>6-8$ arc min, we found that the grid lay either in the fixation plane [as expected according to Fig. 1C(1)], or moved forward to the depth expected of the oblique matching in Fig. 1C(2). The former was seen when the sum of the displacements, $S_L + S_R$, was small (<1 inter-dot spacing), and the latter when the sum was large (>1.5 inter-dot spacing). For intermediate values of the sum $S_L + S_R$, either depth percept might be seen on a given trial, with a probability depending on the sum $S_L + S_R$.

A radically different type of behaviour was seen when the inter-dot spacing was <6 arc min; it was then possible to obtain depth percepts which could not be accounted for by any discrete matching of points. For example, Fig. 3A shows a depth profile of a grid with $S = 0.2$, $S_R = 0.8$. This profile is slanted in depth: it runs continuously between the depths which would be expected from matching the end points (open circles), and none of the intermediate values is compatible with discrete pairings of points within the grid. It is as though the end points have been paired and the intermediate depths obtained by interpolation between these extremes. For comparison, Fig. 3A also shows profiles for $S_L = S_R = 0$ and $S_L = S_R = 0.67$; both of these are

Fig. 1 A, The grid of our stereogram, showing the test dot. B, A single row from the grid, with the left-most dot in the left eye displaced inwards by a fraction S_L , and the right-most dot of the right eye displaced by a fraction S_R of the inter-dot spacing. C, (1), (2), Two possible matchings of points of the left and right eyes in a single row. In (1) the dots are matched in sequence; in (2) the displaced dots at the ends of each eye's row are unmatched, and a point in the left eye is matched to a point in the right eye which is one step of the grid to its left. D, The depth percepts expected, based on conventional stereoscopic theory, from the matchings (1) and (2). The fixation plane (vertical matching in C) is indicated by an arrow. In fact, a fixation target was always presented before the grid was shown, but was never shown simultaneously with it. In (1), two dots lie in front of the fixation plane. Here S_L is shown smaller than S_R , so the right-most dot should be farther in front than the left-most. In (2) the grid lies in front of the fixation plane (corresponding to a crossed disparity). The two monocular points (open circles) appear to lie behind the grid, although we do not know precisely what perceived depth should be assigned to them.



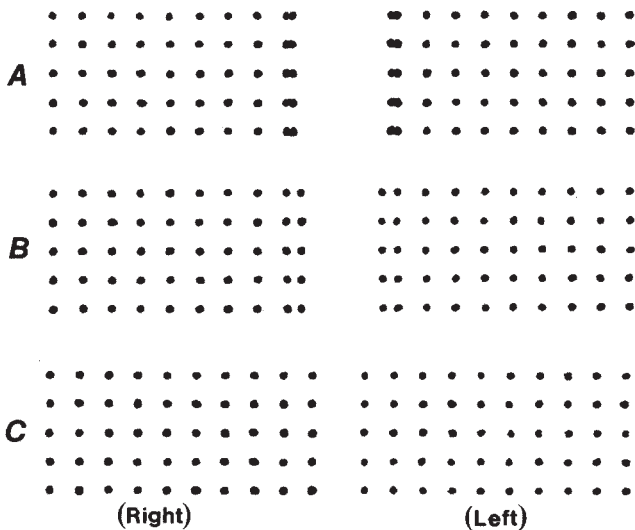


Fig. 2 A stereogram of grids (with dimensions somewhat different from those in Fig. 1A) to show the basic phenomena. The displacements S_L and S_R are equal: they have a large value in panel A and are zero in C. In B, a displacement has been chosen which should give an intermediate depth percept under interpolation conditions. Those who free-fuse by converging will see the grid correctly; those who diverge will see the depths reversed. When the stereogram is viewed quite close up (say 1 m), the discrete matchings of Fig. 1C and D should be seen. When it is viewed from a distance of ≥ 3 m, the inter-dot spacing comes within the interpolation range (< 6 arc min), and B should appear at a depth between that of A and C.

flat, the former being close to the fixation plane, and the latter forward of it, though not as far forward as would be expected for the oblique matching.

These intermediate-depth percepts do not correspond to an ambiguous alternation between the two matchings, as demonstrated by the curves showing the proportion of trials in which the test dot lay in front of the grid, as a function of test dot disparity (Fig. 3B). If there were an alternation of percepts, these curves would be broad, encompassing the range from the fixation plane to the forward matching. However, they are very narrow, having a standard deviation of 10–15 arc s. The intermediate depths (such as curves b–d in Fig. 3B) are, in fact, as well localized as the extremes (curves a, e) which satisfy a 'conventional' matching. Depth percepts which lie between possible discrete matches have been described by Foley³, and may be related to those we find, but his conditions were very different from ours, and the intermediate states far less localized.

Interpolation appears to be quite stable under prolonged inspection (for example, Fig. 2 viewed at ≥ 3 m), and, in particular, the percept does not depend on vergence. The viewing time of 160 ms used in our experiments was designed to eliminate changes of vergence during each presentation. In one experimental condition, the displacements S were changed randomly from trial to trial to prevent the observer from maintaining a biased convergence position in front of the fixation plane, and essentially similar results were obtained.

One might attempt to account for interpolation by saying that the stereo matching system has a coarser resolution than 6 arc min, and 'sees' only the ends of the grid when the inter-dot spacing is small. We can show that this is not the case by randomly perturbing points in the periodic region by the same amount in each eye; when this is done, the grid can be matched as a flat surface in the fixation plane (Fig. 4A), but the oblique matching should introduce disparity variations (Fig. 4B). There are general arguments for expecting that the matching which gives the more continuous depth profile will be preferred⁴ and, indeed, even quite small random perturbations keep the grid in the fixation plane matching Fig. 4A, no matter how large the displacements S . For example, for an inter-dot spacing of 3.7 arc min, a 20-arc s (SM) or 40-arc s (GM) perturbation serves to

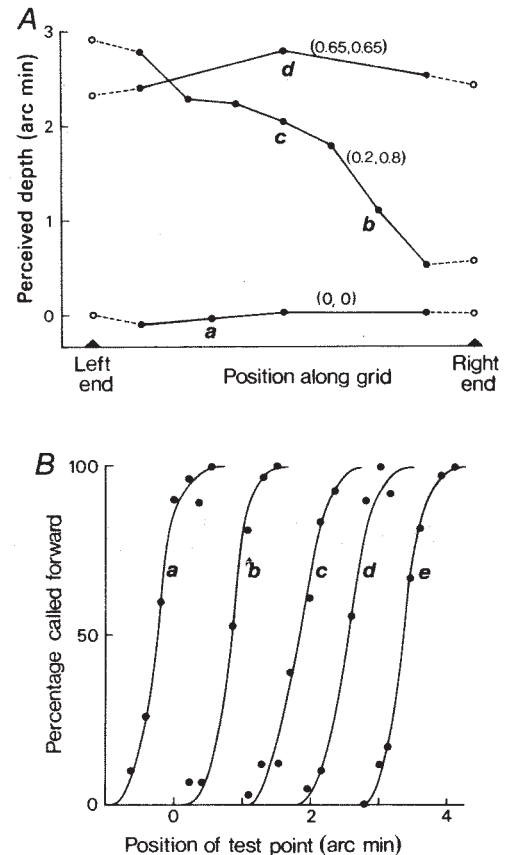


Fig. 3 A, Apparent depths for various values of S_L and S_R (shown in brackets) at selected positions along a grid of inter-dot spacing 3.7 arc min (test subject G.J.M.). Each solid point represents the mean of a psychometric curve (see B) based on at least 200 trials. The open circles show the expected depths of the paired end points of the grid, calculated from their disparities. B, Psychometric functions for various points in A (curves a–d) and for the grid with $S_L = S_R = 1$ (e). The ordinate shows, for a given disparity of the test dot, the percentage of responses 'test dot in front of grid'. If a cumulative gaussian is fitted to these points using the probit method², its mean defines the apparent depth of the grid, and its standard deviation indicates the imprecision of this depth judgement.

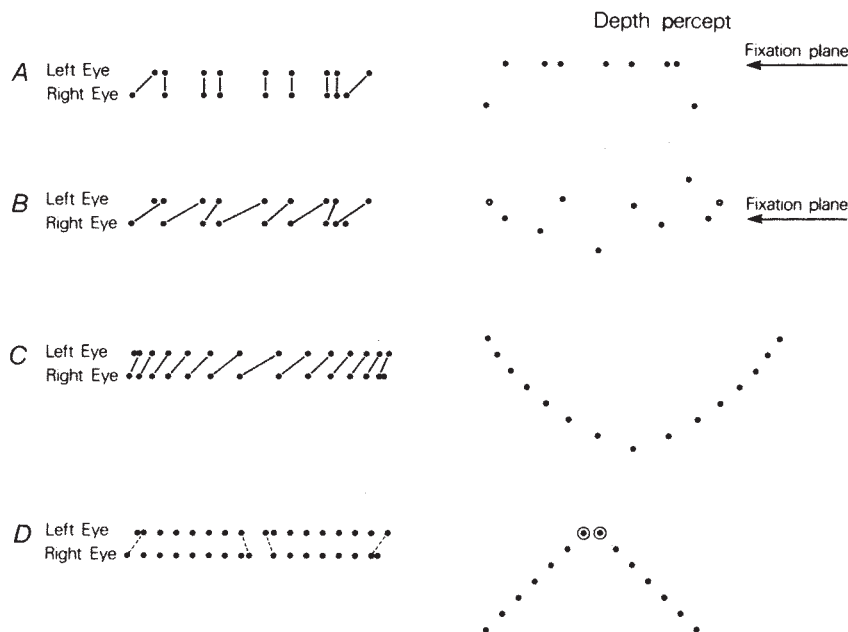
keep the grid in the fixation plane, even with the largest possible values of S_L and S_R . Stereoscopic matching is therefore sensitive to very small perturbations of the points in the grid.

What is the relationship between interpolation and the conventional point-by-point matching seen for larger dot spacings? To determine how they interact, we constructed a grid which expands slowly so that the spacing in the centre is in the range where discrete matching occurs, while the spacing at the edges lies well in the interpolation range (Fig. 4C). A displacement at the edges brings forward the whole grid, generating a vivid arch in depth. This implies that interpolation is not merely an interpretative stage which supplies appropriate depths after matching is complete, but can influence matches, as in the central region here.

We have seen that the edges of the grid have a privileged role in that they provide end points for interpolation. What counts as an edge? If we displace two dots in the centre of our grid, so making a small gap, the resulting percept exhibits two segments of interpolation (Fig. 4D), suggesting that the inner gap provides two 'edges' in this sense (dotted lines).

In summary, the stereo matching system appears to skip over regions with a periodic spacing of < 6 arc min, but is alert to small irregularities within such a region. These irregularities might be interpreted either as gaps or as disparity discontinuities. Of the two possibilities, the second is the more attractive; it implies that it is not periodicity *per se* which is required for

Fig. 4 *A*, In a modified version of the stereogram, a grid is constructed from several identical rows such as that shown here. This row is derived from that shown in Fig. 1*B* by leaving the outer two dots in each eye unchanged, but randomly perturbing the inner dots by the same amount in each eye. The lines show the matching which corresponds to Fig. 1*C*(1), and the depth percept is shown to the right. This is found for all dot spacings in the interpolation range (<6 arc min). *B*, An oblique matching, corresponding to Fig. 1*C*(2), in which each dot in the left eye is matched with a dot one further along in sequence in the right eye; this generates a ragged depth profile (shown on right). *C*, A modified version of the row shown in Fig. 1*B*. The spacing between dots at the edges is 3 arc min; this spacing increases gradually towards the centre, reaching a maximum of 11 arc min, and then decreases again, in a symmetrical fashion. The expected depth percept for the oblique matching shown is a forward-bending arch. *D*, A gap is created by displacing two dots near to the centre of the grid. In the left eye, a dot is moved half an inter-dot spacing to the right; in the right eye, the dot which would be paired with this in the fixation plane is moved to the left by the same amount. With the dot spacing in the interpolation range, two segments of interpolation are seen (shown on right). The fine broken lines indicate the matching which defines the 'edge'. In fact, the perceived depth of these dots is somewhat variable between trials, and this is indicated by ringing of the dots in the diagram of the depth percept.



interpolation, but simply the absence of disparity discontinuities. There is an obvious economy in making explicit matches only where a potential disparity discontinuity is registered.

These findings impose some constraints on models of stereoscopic matching. For example, the fact that disparities at the edges determine the matching within each grid in the absence of vergence movements argues against the coarse-to-fine vergence strategy postulated by Marr and Poggio⁵. Information might, in principle, be propagated inwards from the edges by a cooperative process of point-by-point matching^{4,6}, but it seems unlikely that such a mechanism operates under interpolation conditions.

We thank Professor H. B. Barlow and Dr V. S. Ramachandran for helpful comments. This research was sponsored by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under grant AFOSR-82-0345 and by the Smith-Kettlewell Eye Research Foundation.

Received 17 August 1984; accepted 22 March 1985.

1. Julesz, B. & Chang, J. J. *Biol. Cybernet.* **22**, 107-120 (1976).
2. Finney, D. J. *Probit Analysis: A Statistical Treatment of the Sigmoid Response Curve* (Cambridge University Press, 1952).
3. Foley, J. M. *Vision Res.* **16**, 1263-1267 (1976).
4. Marr, D. & Poggio, T. *Science* **194**, 283-287 (1976).
5. Marr, D. & Poggio, T. *Proc. R. Soc. B* **204**, 301-328 (1979).
6. Julesz, B. *Foundations of Cyclopean Perception* (University of Chicago Press, 1971).

Changes in the dendritic branching of adult mammalian neurones revealed by repeated imaging *in situ*

Dale Purves & Robert D. Hadley

Department of Anatomy and Neurobiology, Washington University School of Medicine, 660 South Euclid Avenue, St Louis, Missouri 63110, USA

A major obstacle to understanding the mechanism of long-term change in the vertebrate nervous system has been the inability to observe the same nerve cell at different times during the life of an animal. The possibility that changes in neural connectivity underlie the remarkable flexibility of the nervous systems of mammals has therefore not been tested by direct observation. Here, we report studies in which we have visualized the same neurone in the superior cervical ganglion of young adult mice at intervals of up to 33 days. This collection of nerve cells is particularly accessible and therefore well suited to our approach. We find that the dendritic branches of the neurones examined change appreciably over intervals of 2 weeks or more; some branches retract, others elongate and others seem to form *de novo*. The apparent remodelling of these postsynaptic elements implies that the synaptic connections of these cells normally undergo significant rearrangement beyond what is usually considered to be the developmental period.

Sexually mature male mice (8-12 weeks old, 25-35 g, CF1 strain) were anaesthetized with chloral hydrate and mounted on a platform that could be moved in three axes for positioning and focusing. Either the right or the left superior cervical ganglion was exposed through a midline incision, care being taken to preserve the blood supply. The wound was superfused with sterile lactated Ringer's solution (Travenol) and the pool of fluid grafted with a chloride-treated silver wire. The ganglion surface was illuminated at an oblique angle by a plastic light guide (diameter 0.7 mm). Proper adjustment of the angle of illumination caused the neurones on the surface of the ganglion to appear as a mosaic of polygonal cell outlines (Fig. 1*a*). Individual neurones could be identified as a result of variations in size, shape, neighbour relationships and position with respect to blood vessels. A particular nerve cell was selected for study and 35-mm colour transparencies were taken of the relevant region of the ganglionic surface at a magnification of $\times 35$ -320. Following photography, the neurone that had been selected was impaled under direct vision at a magnification of $\times 125$ with a microelectrode filled with a 6% solution of 5(6)-carboxyfluorescein in 0.44 M KOH (pH 7.0; resistance 80-150 M Ω). Resting potentials usually ranged from -40 to -60 mV; all cells had overshooting action potentials, occurring either spontaneously or elicited by current injection through the microelectrode. The illuminator was then turned off and the dye injected iontophoretically by passing hyperpolarizing current for 2 min (3-4 nA d.c.). Subsequent observations of the cell were made on a video monitor with a low-light level video camera (GE model 4TE56 SIT) and a $\times 40$ water-immersion lens with a

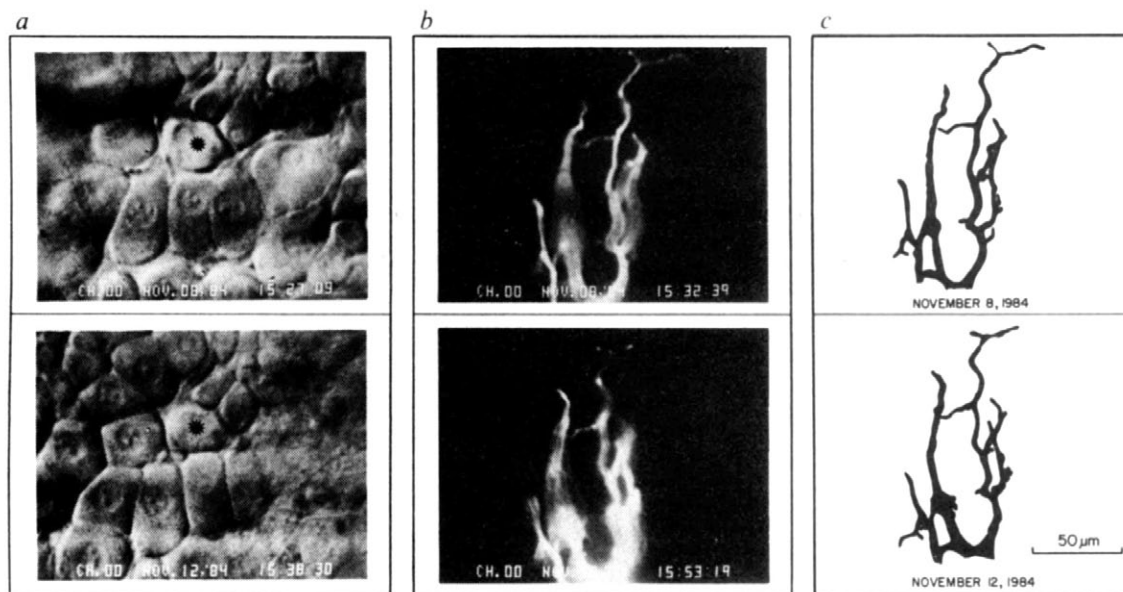


Fig. 1 Identification and imaging of the same neurone over a 4-day interval. *a*, Surface images of a mouse superior cervical ganglion. The same ganglion cells could be recognized at a second operation by virtue of their size, position and relationship to other landmarks such as blood vessels. The neurone marked with an asterisk was impaled with a microelectrode and injected with 5(6)-carboxyfluorescein on both days (see text). *b*, Fluorescent images of a portion of the dendritic arborization of the injected cell in comparable focal planes (this and subsequent photographs were taken from a video monitor). Parts of several large dendrites and their higher-order branches can be seen; the cell soma is just below the bottom of the frame. *c*, Five to ten images at slightly different planes-of-focus were combined to form reconstructions such as those shown here. Comparison of the initial and final images of this and other neurones observed at intervals of a few days showed relatively little change (see Table 1).

numerical aperture of 0.75 and a working distance of 1.6 mm. The dye-filled cell was imaged using standard epifluorescence optics (390–490 nm excitation, 515 nm suppression); low-level illumination was provided by a 6-V, 15-W tungsten bulb controlled by a rheostat and focused through a heat filter. After verifying that the intended cell had been successfully injected, a portion of its dendritic arborization, usually one primary dendrite and its branches, was chosen for study. Fluorescent images were recorded on Polaroid film (Fig. 1*b*). We photographed only a part of the dendritic arborization of each cell to minimize the time of exposure of the cell to the exciting wavelengths. The total period of exposure of the neurone to the exciting light was <1 min. After satisfactory pictures had been obtained the wound was closed and the animal was allowed to recover. The entire procedure usually took ~1 h.

Following an interval of 3–33 days, the animals were re-anaesthetized and the ganglion exposed as before. Projection of the transparencies made at the initial operation allowed us to identify the neurone that had been impaled previously. The cell was re-impaled and injected with dye in exactly the same manner as at the first operation (the dye is lost from the cell within a few hours). Comparisons of dendritic geometry were made either from single focal planes (Fig. 1*b*) or by reconstructing the relevant portion of the dendritic arborization from 5–10 photographs taken at slightly different depths of focus (Fig. 1*c*).

Single neurones from a total of 57 mice were examined by this method. In 44 of these animals we were successful at the second operation in identifying the portion of the ganglionic surface that had been studied previously. In the other 13 animals we were unable to locate the appropriate region; most of these failures occurred near the beginning of our series of experiments when we were less experienced at choosing locations and cell groupings that would be easy to identify subsequently. In 28 of the 44 animals in which the correct region was located, the cell of interest was identified and successfully re-impaled. Average resting potential and action potential amplitudes were not appreciably different from values obtained at the initial operation. In the 16 remaining cases the neighbouring cells could be identified but the cell of interest apparently had died and been removed by phagocytic action. Thus the survival rate of neurones impaled and imaged in this way was ~64%. The

death of about a third of the neurones is not explained by illumination of the fluorescent dye, because in additional control experiments 4 out of 10 cells that were filled with dye but never illuminated also died. The fluorescent dye itself has been shown to be non-toxic in at least one other neural system¹, although prolonged or intense exposure to the exciting wavelength can kill dye-filled cells². Therefore, it seems likely that impalement by a microelectrode is sufficient to cause the death of some cells.

The dendritic arborizations of neurones observed in this way showed shape changes that were, in general, proportional to the

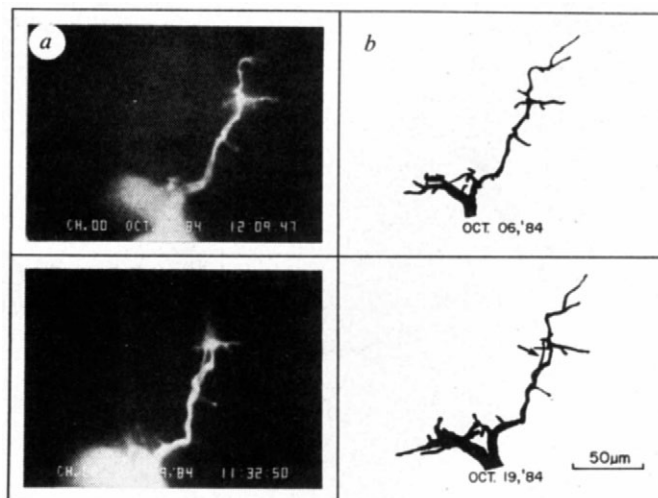


Fig. 2 Images of a neurone observed at an interval of 13 days. *a*, Fluorescent images of an identified neurone at the beginning and end of the period of observation; the plane of focus is approximately the same in both images. *b*, Reconstructions of this portion of the dendritic arborization of the neurone from six different focal planes at the beginning and end of the interval. Examination of higher-order branches showed changes that were more obvious than those observed at shorter intervals; for example, the branch marked by the arrow seems to have formed *de novo* about halfway along the major dendrite.

Table 1 Dendritic changes of mouse superior cervical ganglion cells over intervals of 3–33 days

Interval (days)	No. of cells studied	No. of branches unchanged	No. of branches extended	No. of branches retracted	Branches changed (%)
3–8	4	17	0	0	0
13–14	3	11	4	3	39
28–33	6	9	8	10	67

The classification of a branch as having extended, retracted or remained unchanged is based on an apparent change of $>10\text{ }\mu\text{m}$; measurements were made from camera lucida drawings (see Figs 1–3) using a digitizing pad attached to a computer. Only terminal dendritic branches reconstructed completely from multiple focal planes were included (13 of the 28 cells imaged at two different times).

interval between observations. Although the images obtained at the two different times were never exactly congruent (in part because of small differences in rotation and variations in the tension on the cell), the general configuration of the processes that had been viewed initially was easily recognized and always had the same overall arrangement. Thus, dye filling was complete, or at least consistent in what it revealed. Detailed examination showed differences in the dendritic geometry of many cells, involving the retraction of some branches and the extension of others. In some instances branches evident during the initial observation of the cell had disappeared altogether, whereas other dendritic branches appeared *de novo* in this interval. In most cases, the changes observed at longer intervals were more extensive than those observed at shorter intervals.

Thus, eight neurones that were visualized at intervals of 3–8 days showed relatively minor changes (Fig. 1). The 13 neurones that were imaged after 13–16 days often showed more obvious changes (Fig. 2). In seven animals the same cell was imaged after an interval of 28–33 days. Although the general configuration of the dendrites at the second viewing could still be recognized as similar to the initial pattern, in each case the changes were more extensive than those observed at the shorter intervals (Fig. 3; Table 1). Complete reconstructions of the relevant portion of the dendritic arborization were made in 13 of the 28 cells studied. To quantitate this progressive change, individual branches were measured and scored according to whether they had changed by more than $10\text{ }\mu\text{m}$ (see Table 1).

Our results show that the dendrites of the cells observed gradually change over periods of a few weeks. We cannot eliminate the possibility that such changes represent a response to the experimental procedure rather than normal remodelling of dendritic arborizations. However, two aspects of our results

provide evidence against artificially induced changes in cell shape. First, the surviving cells remained healthy; resting potential and action potential amplitudes of the cells at the second observation were indistinguishable, on average, from values obtained during the initial impalement, and the dendritic arborizations, although changed, were normal in appearance. Second, and perhaps more importantly, the dendritic arborizations changed progressively over the period of observation (up to 33 days) (Table 1). Although the procedure used here might have caused some initial damage, it seems unlikely that such an effect would continue to contribute to the loss and gain of dendritic processes weeks later.

Our interpretation of these observations is that extension and retraction of dendritic processes is a normal occurrence in sympathetic ganglia of young adult mammals. As most synapses on superior cervical ganglion cells in the mouse are made on dendritic processes (ref. 3 and C. J. Forehand, unpublished observations), the dendritic changes we describe are presumably accompanied by commensurate changes in the synaptic connections of these neurones. It will be of interest to examine neurones in other parts of the nervous system where this same general approach may be feasible.

This work was supported by the USPHS and the Muscular Dystrophy Association. We thank J. Voyvodic for help with the computer analysis and C. Forehand, J. Lichtman, J. Sanes, W. Snider, E. Rubin and J. Voyvodic for comments on the manuscript.

Received 7 December 1984; accepted 12 March 1985.

1. Kater, S. B. & Hadley, R. D. *Cytochemical Methods in Neuroanatomy* (eds Palay, S. & Chan-Palay, V.) 441–459 (Liss, New York, 1982).
2. Miller, J. P. & Selverston, A. I. *Science* **209**, 702–704 (1979).
3. Purves, D., Forehand, C. J. & Lichtman, J. W. *Soc. Neurosci. Abstr.* **10**, 1082 (1984).

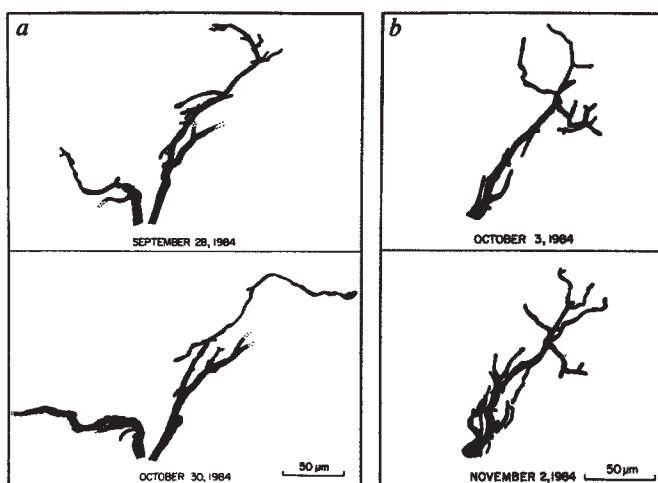


Fig. 3 Reconstruction of a portion of the dendritic arborizations of two ganglion cells observed after an interval of 1 month. *a*, A dendrite with several branches has withdrawn, whereas another branch has apparently extended more than $50\text{ }\mu\text{m}$. Dashed lines indicate continuance of some branches to deeper levels that were not photographed. *b*, Similar retraction and extension of particular branches observed after 28–33 days were more prevalent and extensive than those observed at shorter intervals (compare Figs 1, 2; see also Table 1).

Axon guidance in cultured epithelial fragments of *Drosophila* wing

Seth S. Blair, Marjorie A. Murray & John Palka

Department of Zoology, University of Washington, Seattle, Washington 98195, USA

Growing axons can be guided by a number of different cues: adhesive substrates^{1,2}, diffusible factors^{1–3}, electrical fields^{1,2,4} and even factors intrinsic to the neurone itself^{1,2,5} have all been shown to affect axon orientation and outgrowth *in vitro*. However, in most intact systems it has proved difficult to test directly the role played by these putative guidance cues. Here, we describe a system, the developing wing of the fruitfly, in which we have tested simultaneously two putative guidance mechanisms, physical constraints to axon growth (channels) and the position of neuronal somata (guideposts), using surgical techniques. We show that pioneer sensory axons can navigate correctly and form their normal stereotyped pattern of axon bundles in wing fragments that apparently lack both physical and neural cues. This technique allows access to the surface along which neuronal pathfinding takes place, making possible a wide range of experimental manipulations on the developing system.

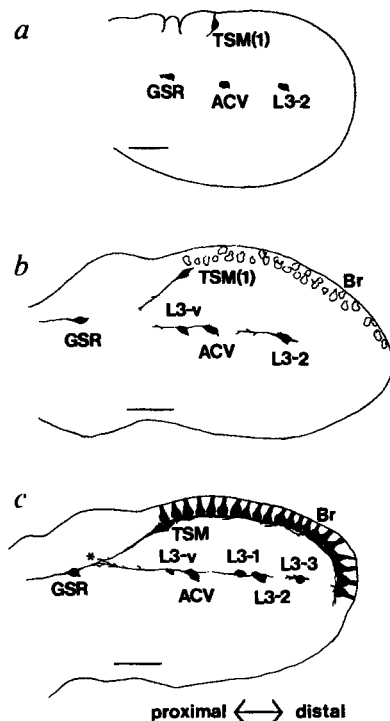


Fig. 1 Sensory neurons of a wing at 2 h after pupariation (a), or after having been removed at 2 h after pupariation and cultured for an additional 8 h (b) or 22 h (c). The earliest-arising neurones are those of the anterior cross vein campaniform sensillum (ACV), the second dorsal campaniform sensillum of the third vein (L3-2), the first twin campaniform sensillum (TSM(1)) and the giant campaniform sensillum of the radius (GSR); these neurones initiate axonogenesis just before (TSM(1) and GSR) or just after (ACV and L3-2) 2 h after pupariation. The ACV, L3-2 and TSM(1) neurones are each accompanied by an 'extra' neurone, E-1, E-2 and E-T respectively, which seem to develop similarly to the sensory neurones⁸; each of these cell pairs will be treated as a single unit designated by the name of the sensillum. Later-arising neurones initiate outgrowth at about the same times *in vivo* and *in vitro*, the ventral campaniform sensillum (L3-v) neurone at 5 h after pupariation, and the neurones of the first and third dorsal campaniform sensilla of the third vein (L3-1 and L3-3) and of the bristle sensilla (Br) at 12 h after pupariation^{8,14}. Axon outgrowth is slightly slowed but essentially normal *in vitro*¹⁴, although axons often fail to grow proximally beyond the junction of the first and third veins (asterisk; see text) after 22 h *in vitro*. Wings were cultured in Shields and Sang M3 medium supplemented with 10% fetal bovine serum²⁵ as described previously¹⁴, except for the addition of 250 U ml⁻¹ penicillin, 250 µg ml⁻¹ streptomycin, 0.63 µg ml⁻¹ Fungizone (Gibco) and 50 µg ml⁻¹ tetracycline-HCl (Sigma), and the exclusion of β-ecdysone from the medium. Distal is to the right. Scale bars, 50 µm.

The wing of the fruitfly *Drosophila melanogaster* develops from an epithelial sac, the wing imaginal disk. During early stages of metamorphosis, the disk everts and flattens, bringing together originally separate epithelia to form the dorsal and ventral surfaces of the flat wing blade. Gaps are formed between the apposed inner epithelial surfaces, so that by 4 h after pupariation a system of channels loosely corresponding to the pattern of venation of the adult wing is created⁶⁻⁸. These channels are transiently lost when the wing inflates between 12 and 24 h after pupariation^{6,7}. The pioneer sensory neurones of the wing (the neurones of specific large campaniform sensilla) are born (before 0 h after pupariation, from epithelial mother cells), differentiate (0–2 h after pupariation), initiate axonal outgrowth (1–3 h after pupariation) and complete the basic scaffolding of nerve bundles (12 h after pupariation) before this time^{8,9} (Fig. 1). Thus, the first axons navigate through a system of channels between apposed epithelial cells, which might act to limit and guide axonal outgrowth. A similar role has been proposed for

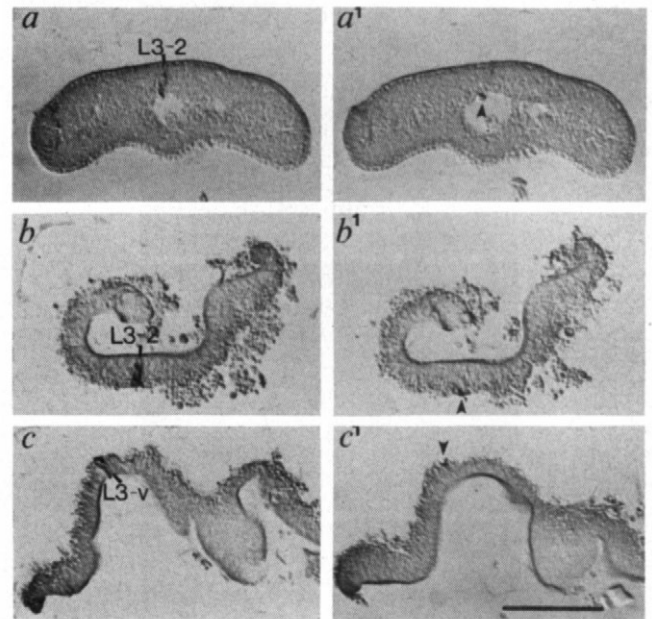


Fig. 2 Transverse sections through wings that had been removed at 2 h after pupariation and cultured an additional 8 h. Neurones were stained with anti-horseradish peroxidase (HRP) and an HRP-conjugated second antibody, and appear dark because of the HRP reaction product. a, a', Adjacent sections from a whole cultured wing, showing channels and the position of the soma of the L3-2 neurone (a) and axon (a', arrow) within the most central of them. b, b', Adjacent sections from a dorsal fragment generated at 2 h after pupariation, showing the position of the L3-2 soma (b) and axon (b', arrow); note the lack of channels or folds along the path of axonal outgrowth. c, c', Adjacent sections from a ventral fragment, showing the position of the soma of the L3-v neurone (c) and short axon (c', arrowhead); this fragment similarly lacks channels or folds along the path of axonal outgrowth. Three or more of each type of fragment (dorsal and ventral) have been completely sectioned; none shows any type of morphological irregularity which might act to guide growing axons in the distal wing blade. Fragmented wings for sectioning were reared and fixed after having been pinned flat onto Sylgard using glass needles; this reduces curling of the wings and does not affect axon guidance (8 h dorsal fragments, 6 of 6 cases). Wings were fixed and stained with anti-HRP as described previously⁸, except that HRP-labelled goat anti-rabbit IgG (Cappel) was used as the second antibody, reacted with diaminobenzidine, cleared in 80% glycerol, embedded in JB-4 plastic, sectioned at 5 µm (in some cases (not shown) counterstained with toluidine blue) and viewed using Nomarski optics. Scale bar, 50 µm.

channels in the developing vertebrate central nervous system¹⁰ and retina¹¹.

To test this hypothesis, the formation of channels between the wing epithelia was prevented by cutting the wing disk into dorsal and ventral fragments, after eversion but before the apposition of the two surfaces (2 h after pupariation); channels in such fragments could no longer be closed because the opposite epithelium was lacking. Axon outgrowth was assessed by rearing the fragments in culture for a further 8 or 22 h and then staining their neurones with antiserum generated against horseradish peroxidase¹². No closed channels were observed in epithelial fragments cultured for 8 h; axons grew along the innermost surface of the epithelium and did not seem to be following any histologically detectable morphological irregularities, such as grooves or folds (Fig. 2). This technique therefore allowed us to analyse the pathfinding abilities of axons growing in a channel-free environment.

As demonstrated previously^{13,14}, whole wing disks reared *in vitro* show largely normal patterns of axonal outgrowth (Fig. 1). Neurones in dorsal and ventral wing fragments differed somewhat in their ability to reproduce this normal pattern¹⁵.

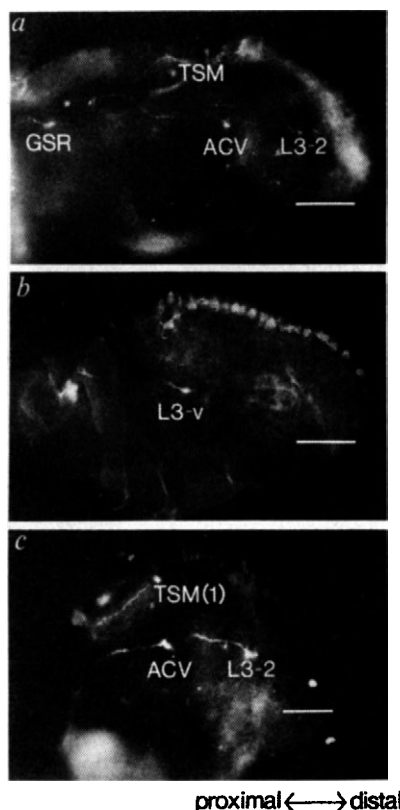


Fig. 3 Cultured wing fragments stained with anti-HRP. *a*, Dorsal wing fragment raised *in vitro* for 22 h. The nerve bundles of the first and third veins have formed normally and join at approximately their normal site just distal to the GSR neurone (compare with Fig. 1), although in this case they failed to grow to the GSR neurone (asterisk marks gap; see text). *b*, Ventral wing fragment, raised *in vitro* for 22 h. The L3-v neurone has formed a short bipolar projection along the proximo-distal axis of the wing; the proximal process is longer than the distal. *c*, Dorso-distal fragment, raised *in vitro* for 8 h. Although the fragment lacked both the GSR neurone and closed channels, the remaining ACV, L3-2 and TSM(1) neurones grew along the correct paths and directions (compare with Fig. 1). Fragments were fixed and stained as described previously¹⁴. Distal is to the right. Scale bars, 50 μ m.

The earliest-developing neurones of the distal wing blade (ACV, TSM(1) and L3-2; see Fig. 1 legend) are all found on the dorsal surface. In dorsal fragments examined after 8 h in culture, the axons of these early neurones had grown in the correct direction along the normal paths (Table 1), either proximally (ACV and L3-2) or proximoposteriorly (TSM(1)). In fragments raised for 22 h, these neurones, together with the later-arising L3-1, L3-3 and Br neurones, formed largely normal nerve bundles (49 of 50 cases; Fig. 3*a*). The only commonly observed abnormality was that the ACV and TSM(1) axons failed to reach the GSR neurone (33 of 49 cases), even though they contacted each other at the usual, slightly more distal, point. This is also common in whole wings raised *in vitro*¹⁴ (under current conditions, 19 of 33 cases) and probably results from damage to the wings or suboptimal culture conditions. In any case, initial outgrowth and most other elements of nerve formation were normal in dorsal fragments despite the lack of channels.

In contrast, axonal outgrowth of the only ventral neurone of the distal wing blade was both delayed and variable in the absence of the dorsal surface. In whole wings, this neurone, the L3-v, initiates axonogenesis only after the early dorsal axons have grown past it (~5 h after pupariation); its process then joins with and apparently follows the early axons. Although this fasciculation obscures the details of further L3-v outgrowth, dye fills in adults demonstrate that the L3-v forms a single, proximally directed axon¹⁶. In ventral fragments reared for 8 h,

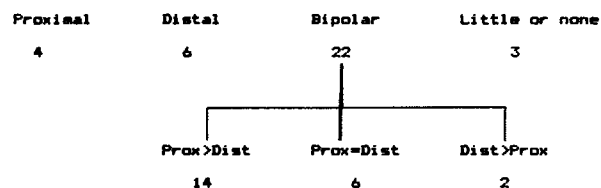


Fig. 4 Direction of outgrowth of the axon of the L3-v neurone in ventral fragments, generated at 2 h after pupariation and cultured for an additional 22 h. Significant outgrowth was taken as outgrowth longer than one neuronal cell diameter. Some branching of axons was observed, but in all cases outgrowth was oriented roughly along the proximo-distal axis of the wing. In five cases a region of dorsal epithelium lay along the anterior of the ventral fragment near the L3-v; in these wings (not included) the L3-v axon grew anteriorly into this region and sometimes seemed to contact the short nerve formed by the TSM neurones. Obviously infected cultures were not included in the data pool.

the L3-v formed no or only very short processes (<1 cell diameter, 20 of 20 cases). By 22 h it usually formed a short bipolar projection (Figs 3*b*, 4); in a few cases, exclusively proximal or distal outgrowth was seen. Because the normal growth pattern of this neurone is not known in detail, it is not certain that these results represent abnormal axonogenesis. However, the variable and generally stunted outgrowth seen in ventral fragments suggest that L3-v axonogenesis was to some extent disrupted by the removal of the dorsal epithelium. Nevertheless, important features of L3-v axonogenesis were retained in ventral fragments: outgrowth was still oriented approximately along the proximo-distal axis of the wing, and was usually, though not always, longer in the proximal than in the distal direction (Fig. 4). Thus, correctly directed axon outgrowth can occur even in a fragment in which the neuronal population is reduced to the greatest possible extent, namely a single cell differentiating in an expanse of epithelium.

Thus, it is apparent that both dorsal and ventral fragments must contain some types of guidance cues other than channels. For example, the filopodia of a neurone might seek out a more proximal 'guidepost' or 'landmark' neurone, which would act as a target for axon growth¹⁷⁻¹⁹. Guidepost neurones have been shown to have no necessary role in normal *Drosophila* wings^{14,20} but might effectively guide neurones in dorsal fragments lacking channels. This possibility was tested by generating fragments, at or before the time of initial axon outgrowth (1.5-2 h after pupariation), from which both proximal tissue (containing a putative guidepost neurone) and the ventral epithelium had been removed. When such dorso-distal fragments were cultured for 8 h, the ACV and TSM(1) neurones still navigated along their normal paths and in their normal directions, even though they lacked both channels and the proximal putative guidepost neurone (GSR) towards which they normally grow (Fig. 3*c*, Table 1). Thus, the GSR neurone does not act as a necessary guidepost for these dorsal neurones, even in the simultaneous absence of obvious physical cues.

Table 1 Fraction of neurones showing normal axonal outgrowth in dorsal and dorso-distal wing fragments

Fragment	ACV	L3-2	TSM(1)
Dorsal	17/17	13/14	14/14
Dorso-distal	11/11	8/8	8/8

Fragments were cut at 2 h after pupariation and cultured for an additional 8 h. Not all neurones could be observed in each fragment. Dorso-distal fragments all lacked the GSR neurone. The single instance of abnormal L3-2 outgrowth consisted of a lengthening of its normally short distal process; this abnormality has also been observed in whole wings raised *in vitro*. Obviously infected cultures were not included in the data pool.

Although there is strong evidence that guidepost neurones can direct growing axons in specific instances in developing insects^{18,21,22}, even in these cases initial axon outgrowth seems to be directed, as here, by some other cue(s)^{18,21}. A one-dimensional gradient of neural adhesivity may provide one such cue^{23,24}. However, some features of the axonal pattern in fly wings, such as the angled outgrowth of the TSM(1) axon, are not easily explained by a single gradient, and might require the reading of information in two dimensions. An alternative mechanism that would require less from the neurone would be to possess adhesive cues localized selectively along the normal nerve pathways of the wing¹⁶. Finally, polarity information intrinsic to each neurone⁵ might also guide outgrowth, particularly during its initial stages; extrinsic cues might then be needed only to induce large changes in the trajectory of the growing axon. As the surface along which the neurones of the fly wing navigate can now be exposed without disrupting outgrowth significantly, it may be possible to manipulate this surface directly, to study a variety of putative guidance mechanisms.

We thank L. M. Nagy, Dr M. Schubiger and E. Cole. This work was supported by PHS postdoctoral fellowships NS 07185-02 (S.S.B.) and NS 07268-01 (M.A.M.), NIH research grant NS 07778 (J.P.) and NSF research grant BNS 83-19890 (J.P.).

Received 10 December 1984; accepted 1 March 1985.

- Johnston, R. N. & Wessells, N. K. *Curr. Topics dev. Biol.* **16**, 165-206 (1980).
- Letourneau, P. C. in *Neuronal Development* (ed. Spitzer, N. C.) 213-254 (Plenum, New York, 1982).
- Gundersen, R. W. & Barrett, J. N. *J. Cell Biol.* **87**, 546-554 (1980).
- Patel, N. & Poo, M.-M. *J. Neurosci.* **2**, 483-496 (1982).
- Solomon, F. *Cell* **16**, 165-169 (1979).
- Waddington, C. H. *J. Genet.* **41**, 75-139 (1940).
- Anderson, H. *Wilhelm Roux Arch. dev. Biol.* **193**, 226-233 (1984).
- Murray, M. A., Schubiger, M. & Palka, J. *Dev. Biol.* **104**, 259-273 (1984).
- Jan, Y. N., Ghysen, A. & Jan, L. Y. *Soc. Neurosci. Abstr.* **9**, 302.3 (1983).
- Singer, M., Nordlander, R. H. & Egar, M. *J. comp. Neurol.* **185**, 1-22 (1979).
- Silver, J. & Sidman, R. L. *J. comp. Neurol.* **189**, 101-111 (1980).
- Jan, L. Y. & Jan, Y. N. *Proc. natn. Acad. Sci. U.S.A.* **79**, 2700-2704 (1982).
- Edwards, J. S., Milner, M. J. & Chen, S. W. *Wilhelm Roux Arch. dev. Biol.* **185**, 59-77 (1978).
- Blair, S. S. & Palka, J. *Dev. Biol.* **108**, 411-419 (1985).
- Blair, S. S., Nagy, L. M. & Palka, J. *Soc. Neurosci. Abstr.* **10**, 42.4 (1984).
- Palka, J., Schubiger, M. & Ellison, R. *Dev. Biol.* **98**, 481-492 (1983).
- Bentley, D. & Keshishian, H. *Trends Neurosci.* **5**, 354-358 (1982).
- Bentley, D. & Caudy, M. *Cold Spring Harb. Symp. quant. Biol.* **48**, 573-585 (1983).
- Taghert, P. H., Bastiani, M. J., Ho, R. K. & Goodman, C. S. *Dev. Biol.* **94**, 391-399 (1982).
- Schubiger, M. & Palka, J. *Dev. Biol.* **108**, 399-410 (1985).
- Berlot, J. & Goodman, C. S. *Science* **223**, 493-496 (1984).
- Goodman, C. S. *et al. Science* **225**, 1271-1279 (1984).
- Nardi, J. B. *Dev. Biol.* **95**, 163-174 (1983).
- Caudy, M. & Bentley, D. *Soc. Neurosci. Abstr.* **10**, 42.6 (1984).
- Shields, G. & Sang, J. H. *Drosoph. Inf. Seru.* **52**, 161 (1977).

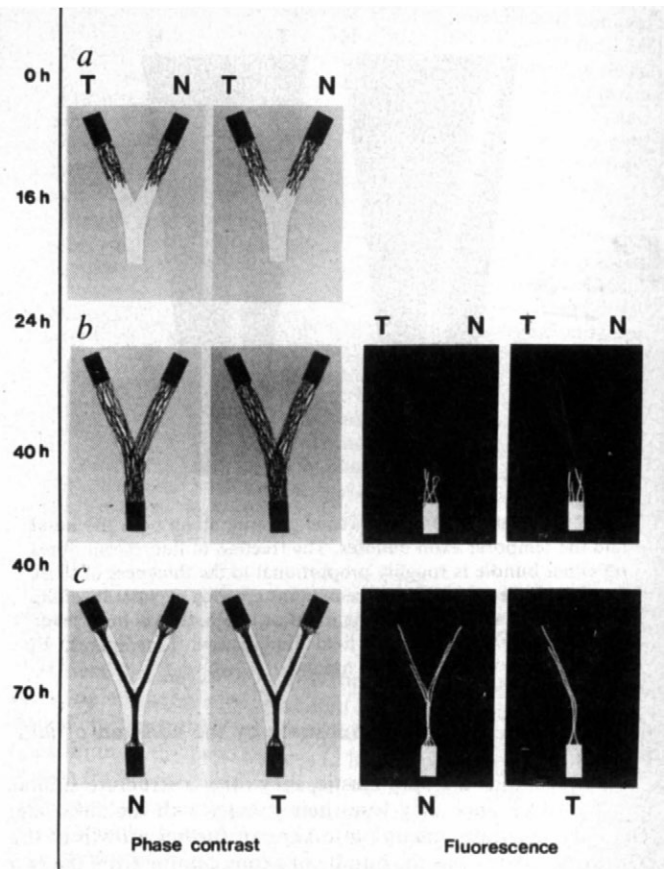


Fig. 1 Schematic description of the experiment. *a*, Phase-contrast diagram of axon growth, with first explantations at 0 h; *b*, explantation of fluorescent explants at 24 h; *c*, addition of α -laminin at 40 h. Axons visualized at 16, 40 and 70 h. T, temporal; N, nasal. **Methods.** Nasal and temporal explants (0.2 $\times$ 2 mm) of chick retina were prepared as described previously⁶. They were placed at time 0 onto laminin-coated round coverslips and incubated in culture medium at 37 °C. Laminin was applied to the coverslips by first treating them with 3-(triethoxysilyl)propylamine and then incubating them in laminin (40 μ g ml⁻¹) for 3 h at 37 °C. The laminin serves as substrate for growing axons which are visible in phase contrast (*a*). To restrict axonal growth to the indicated Y-shaped field, border lines were drawn with a hypodermic needle on the laminin surface. Axons from nasal and temporal explants joined at the stem of the Y after ~20 h. At ~24 h, a fluorescent explant (either temporal or nasal) was placed at the foot of the Y which was reached by the axons a few hours later. At 40 h the Y-shaped field was crowded with axons as seen in phase contrast and a few fluorescent axons growing up the stem of the Y could be seen in fluorescence (*b*). At about that time, anti-laminin (provided by Dr R. Timpl) was added to the culture. After ~1 h, all fibres were detached from the laminin-coated coverslip and the axons spanned Y-shaped bundles joining the explants. Now the further growth of fluorescent axons (which can be inhibited by antibodies against the cell adhesion molecule N-CAM; ref. 7) could only occur along pre-existing axons. As seen in *c*, temporal fluorescent axons strongly prefer to grow along temporal axons, whereas nasal axons grow along both nasal and temporal axons.

Position-dependent properties of retinal axons and their growth cones

Friedrich Bonhoeffer & Julita Huf

Max-Planck-Institut für Entwicklungsbiologie, Spemannstr. 35, D-7400 Tübingen, FRG

The formation of the very orderly neuronal projection from the retina to the optic tectum is not yet understood, but several mechanisms are thought to be involved in a coordinated fashion. These mechanisms may include mechanical or chemical guidance in channels, guidance by spatial gradients of positional markers, gradients of temporal (maturation) markers or specific inter-axon interactions (see ref. 1 for review). The last-mentioned mechanism could explain the fibre order found in optic nerve and tract. It requires that some or all growing retinal axons can distinguish between retinal axons of various origins and grow preferentially along retinal axons originating from the same area as themselves. The *in vitro* experiments described here show that growth cones from the temporal half of the chick retina grow preferentially along temporal axons, whereas growth cones from nasal retina do not distinguish between nasal and temporal axons.

Two 6-day-old chick embryo retinal explants were placed on a laminin-coated coverslip as described in Fig. 1 legend. One explant was taken from the temporal side of the area next to the optic fissure, whereas the other was taken from the nasal side. The axons were forced to grow within predetermined tracks scratched into the laminin so that the axons of both explants met and joined to form a 'Y'-like configuration. (The growth rate of nasal and temporal axons was about the same.) About 24 h later, a third rhodamine-stained fluorescent explant of either nasal or temporal origin was placed near the growth cones at the base of the Y. After a further 16 h, when most of the nasal and temporal axons had reached the labelled explant and some of the fluorescent axons were already leaving it, all axons were



Fig. 2 *a*, Nasal fluorescent axons growing along both the nasal and the temporal axon bundles. The fraction of fluorescent fibres on either bundle is roughly proportional to the thickness of these bundles. *b*, Temporal fluorescent axons growing exclusively along the temporal axon bundle. The nasal axon bundle has been made barely visible by weak dark-field illumination. T, temporal; N, nasal.

detached from the laminin substrate by the addition of anti-laminin.

Because axons are very elastic, they form a structure similar to a tightrope once they lose their contact with the substrate. The only substrate remaining to support further growth of the fluorescent axons was the bundle of axons coming from the two unstained explants; axons had the choice of growing along nasal or temporal fibres within this bundle. Therefore, the preference of the labelled axons became apparent where the bundle bifurcated.

As shown in Figs 1 and 2, all temporal fibres grow towards the temporal explant, that is, they grow preferentially along temporal axons. The results are the same whether the temporal explants stem from more central or more peripheral positions in the retina. On the other hand, nasal fibres show no preference and grow towards either explant. The fraction of nasal fibres growing towards either of the explants is approximately proportional to the relative thickness of the fibre bundle coming from the explant. In Fig. 2*a*, for example, the number of axons growing towards the temporal explant is relatively high because the bundle of temporal fibres was relatively strong.

Of 61 experiments of this type, 53 could be evaluated; in the other 8 cases too few (or no) fibres passed the bifurcation point. The stained explants of the successful experiments were of nasal origin in 34 cases and of temporal origin in 19. Temporal axons grew exclusively towards the temporal explant, whereas nasal axons always grew towards both explants. The fraction of fluorescent axons growing along either unstained bundle was roughly proportional to the relative thickness of these bundles.

According to our results, axons of temporal origin exhibit position-specific properties of self recognition; temporal growth cones recognize temporal axons. From previous observations it is known that temporal axons show a strong tendency to fasciculate with each other (in contrast to nasal axons^{2,3}), in good agreement with the observations reported here. Until now, however, it has not been clear whether this difference in behaviour results from different properties of the nasal and temporal growth cones, or their axons, or both. Our experiments indicate that both axons and growth cones of temporal retinal ganglion cells are different from nasal axons and growth cones.

We have also performed analogous experiments concerning the dorso-ventral axis. Explants were taken from dorsal and ventral positions relatively far apart (half the meridian) from both temporal and nasal retina. These experiments have not revealed any position-specific effects. Anatomical investigations

have shown that chick retinal axons are already sorted out in the dorso-ventral dimension when they arrive at the tectum, whereas no sorting of nasal and temporal fibres has been detected⁴. Therefore, differences between dorsal and ventral axons in our *in vitro* assay might have been expected rather than differences between nasal and temporal axons, but the opposite is in fact the case. Our finding that temporal axons have a strong preference for growing along other temporal axons *in vitro*, however, is in excellent agreement with *in vivo* observations in the frog⁵. In contrast to the nasal fibres, the temporal axons of frog fasciculate with each other in the optic tract and form a dense bundle. Whether and how these position-specific properties of axons and growth cones are used during the formation of the chick retino-tectal projection is not known.

We thank Drs Alfred Gierer and Jonathan Raper for stimulating discussions and helpful comments on the manuscript.

Received 5 December 1984; accepted 20 March 1985.

1. Purves, D. in *Repair and Regeneration of the Nervous System* (ed. Nicholls, J. G.) 107-125 (Springer, Berlin, 1982).
2. Halfter, W., Claviez, M. & Schwarz, U. *Nature* **292**, 67-70 (1981).
3. Bonhoeffer, F. & Huf, J. *EMBO J.* **1**, 427-431 (1982).
4. Thanos, S. & Bonhoeffer, F. *J. comp. Neurol.* **219**, 420-430 (1983).
5. Fawcett, J., Taylor, J., Gaze, M., Grant, Ph. & Hirst, E. M. *J. Embryol. exp. Morph.* (in the press).
6. Bonhoeffer, F. & Huf, J. *Nature* **288**, 162-164 (1980).
7. Edelman, G. M. *Science* **219**, 450 (1983).

Synapsin I is a spectrin-binding protein immunologically related to erythrocyte protein 4.1

Anthony J. Baines & Vann Bennett

Department of Cell Biology and Anatomy, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

The membrane-associated cytoskeleton is considered to be the apparatus by which cells regulate the properties of their plasma membranes, although recent evidence has indicated additional roles for the proteins of this structure, including an involvement in intracellular transport and exocytosis (see refs 1-3 for review). Of the membrane skeletal proteins, to date only spectrin (fodrin) and ankyrin have been purified and characterized from non-erythroid sources. Protein 4.1 in the red cell is a spectrin-binding protein^{4,5} that enhances the binding of spectrin to actin^{4,6,7} and can apparently bind to at least one transmembrane protein⁸⁻¹⁰. Immunoreactive forms of 4.1 have been detected in several cell types, including brain¹¹⁻¹⁴. Here we report the purification of brain 4.1 on the basis of its cross-reactivity with erythrocyte 4.1 and spectrin-binding activity. We further show that brain 4.1 is identical to the synaptic vesicle protein, synapsin I, one of the brain's major substrates for cyclic AMP and Ca²⁺-calmodulin-dependent kinases^{15,16}. Spectrin and synapsin are present in brain homogenates in an ~1:1 molar ratio. Although synapsin I has been implicated in synaptic transmission, no activity has been previously ascribed to it.

Antibodies raised against pig erythrocyte 4.1 recognize only protein 4.1 in pig erythrocyte ghosts (Fig. 1). Cross-reactive polypeptides in brain (rat, pig or calf) have an apparent relative molecular mass (*M_r*) of 75,000 and 73,000 (judged by SDS-gel electrophoresis with a discontinuous buffer system¹⁷).

The cross-reactive polypeptides were extracted from myelin-depleted brain membranes under conditions similar to those used for the extraction of erythrocyte protein 4.1. Thus, they were virtually inextractable in solutions of low ionic strength but were liberated readily from the membranes by treatment with solutions of higher ionic strength (for example >0.2 M KCl). The extracted protein failed to bind DEAE-cellulose (in 10 mM Na-phosphate at pHs as high as 8.2) or hydroxylapatite in the presence of 1 M KCl. However, it was

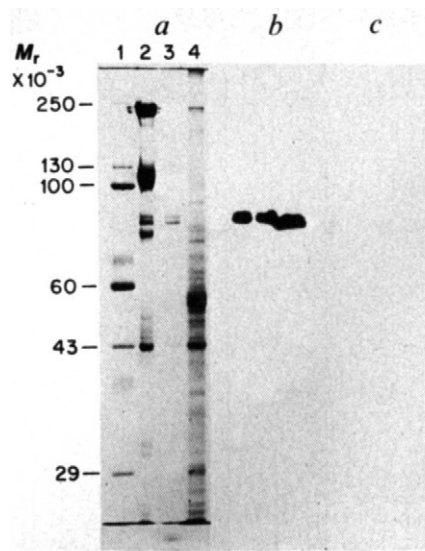


Fig. 1 Identification of a brain protein immunologically related to erythrocyte 4.1. Antibodies were raised in rabbits against SDS-denatured protein 4.1 that had been purified by standard procedures^{4,5} and were affinity-purified on 4.1 immobilized on nitrocellulose strips after SDS-gel electrophoresis and blotting³⁸. In *a*, *b* and *c*, proteins were electrophoresed on 9% SDS-polyacrylamide gels¹⁷ and stained with Coomassie blue (*a*) or transferred to nitrocellulose paper²⁷ (*b*, *c*) for reaction with anti-red blood cell (RBC) 4.1 antibodies and ¹²⁵I-protein A (*b*) or non-immune IgG (*c*). Lanes 1, *M_r* standards: filamin (250,000), vinculin (130,000), α -actinin (100,000), catalase (60,000), actin (43,000), carbonic anhydrase (29,000); lanes 2, pig RBC ghosts; 3, pig RBC 4.1; 4, pig brain post-nuclear supernatant.

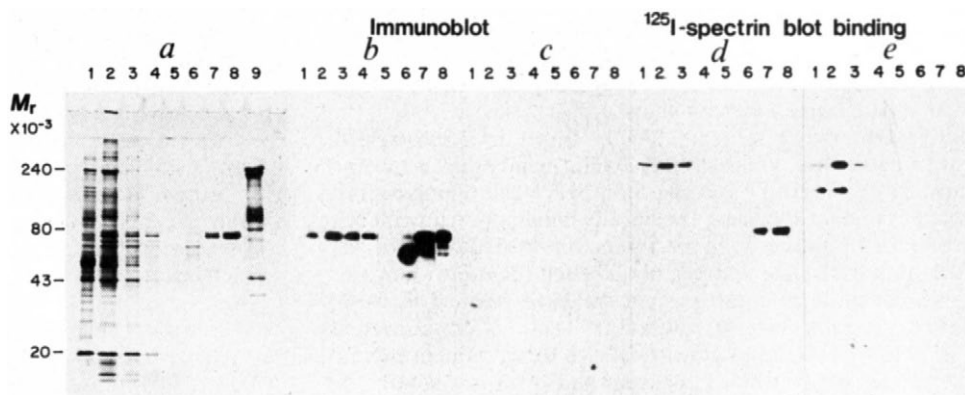
Fig. 2 Purification of brain 4.1. Fractions from the purification procedure were analysed on 5–15% polyacrylamide gels, which were stained with Coomassie blue (*a*) or transferred to nitrocellulose (*b*–*e*). Nitrocellulose replicas were reacted with: *b*, anti-RBC 4.1; or *c*, non-immune IgG followed by ¹²⁵I-protein A. In *d* and *e* they were incubated with ¹²⁵I-RBC spectrin (20 nM, 2×10^6 c.p.m. ml⁻¹ as described in ref. 27) in the absence (*d*) or presence (*e*) of a 20 fold excess of unlabelled spectrin. Lanes 1, myelin-depleted pig brain membranes; 2, 0.25 M KCl extract; 3, first hydroxylapatite (HA) column eluate; 4, DEAE supernatant; 5, carboxymethyl-cellulose (CMC) flow-through; 6, 30 mM NaCl eluate; 7, 60 mM NaCl eluate from CMC; 8, second HA column, MgCl₂ eluate; 9, *M_r* standards (human erythrocyte ghosts). In *b*, lane 8 contains half the amount of protein loaded in the corresponding lanes in *a* and *c*–*e*.

Methods. Myelin-depleted membranes²⁷ from pig brain (350–450 g) were washed twice with 40 mM KCl, 2 mM EGTA in buffer A (10 mM Na-phosphate, 1 mM dithiothreitol (DTT), 1 mM NaN₃, 1 μ g ml⁻¹ leupeptin, 1 μ g ml⁻¹ pepstatin A, 1 μ g ml⁻¹ chymostatin, 200 μ g ml⁻¹ phenylmethylsulphonyl fluoride (PMSF), 1 mM diisopropyl fluorophosphate (DFP) pH 7.4) and were brought to 1:1 with buffer A containing 0.25 M KCl and 2 mM EGTA. Incubation at 37 °C for 30 min released most of the cross-reactive peptides, which were recovered in the supernatant after centrifugation at 30,000g for 30 min. The protein extract was concentrated 10-fold by precipitation with ammonium sulphate at 65% saturation and resuspended in, and dialysed against, 1 M KCl, 1 mM EDTA, 0.1% Tween 20 in buffer A lacking DFP. After centrifugation at 180,000g for 90 min, the supernatant was made 5 mM in MgCl₂ and applied to a 20-ml column of hydroxylapatite (fast flow, Calbiochem) equilibrated with 1 M KCl, 10 mM Na-phosphate, 5 mM MgCl₂, pH 7.4. The void volume, containing the cross-reactive peptides, was dialysed extensively against 1 mM EDTA in buffer A lacking DFP, followed by batchwise absorption with 10 ml DEAE-cellulose (Whatman DE53, equilibrated with buffer A, lacking DFP): cross-reactive peptides were recovered in the supernatant after centrifugation at 1,000g for 2 min. The supernatant was applied to a 10-ml column of carboxymethyl-cellulose (Whatman CM32), equilibrated with buffer B (10 mM Na-phosphate, 1 mM EDTA, 0.1% Tween 20, 0.5 mM DTT, 1 mM NaN₃, pH 7.4), then the column was washed with buffer B and eluted stepwise with 30 mM NaCl in buffer B and 60 mM NaCl, 5 μ g ml⁻¹ leupeptin in buffer B. Peak fractions from the second step were dialysed against buffer C (1 mM Na-PIPES, 0.1 mM DTT, 0.05% Tween 20, 1 mM NaN₃, 1 μ g ml⁻¹ leupeptin pH 6.8) and applied to a 2-ml column of hydroxylapatite (high-resolution, Calbiochem), equilibrated with buffer C. Cross-reactive spectrin-binding peptides were eluted with 5 mM MgCl₂ in buffer C. Yields of 2–5 mg were typically obtained. The protein was typically 90–95% pure when prepared from pig brain, but a contaminant of *M_r* 48,000 was present as 10–15% of the total protein when this method was applied to calf brain. This yield represents $\geq 2\%$ of the total brain 4.1 in a homogenate, major losses occurring because of incomplete extraction from membranes, proteolysis and adsorption to surfaces.

retained by carboxymethyl-cellulose and by hydroxylapatite at low ionic strength and pH 6.8. These properties are those of a very basic protein¹⁸ and allowed the scheme for purification shown in Fig. 2. The purified brain 4.1 displayed anomalous migration between different SDS gel systems. Using a continuous gel buffer¹⁹ we obtained subunits of *M_r* 88,000 and 85,000: in this system, brain 4.1 migrates more slowly than erythrocyte 4.1, as has been observed previously¹².

The spectrin-binding activity of brain 4.1 was shown using 'blot binding' techniques in which fractions were taken from various stages of purification of brain 4.1 and were electrophoretically separated, transferred to nitrocellulose and incubated with ¹²⁵I-labelled erythrocyte spectrin (Fig. 2). As a control for nonspecific binding, additional unlabelled spectrin was included in some of the incubations to displace specific binding. Brain 4.1 at later stages of the purification specifically binds ¹²⁵I-spectrin on the nitrocellulose (Fig. 2*a,b*), but this method was not sensitive enough to detect brain 4.1 in early stages of the purification. Other basic proteins (notably myelin basic protein) do not bind ¹²⁵I-spectrin.

Conversely, ¹²⁵I-brain 4.1 was incubated with blots on which erythrocyte and brain spectrins were immobilized (Fig. 3). As controls for nonspecific binding, we used heat-denatured ¹²⁵I-brain 4.1. As additional internal controls for binding of a basic protein (brain 4.1) to an acidic protein (spectrin), actin, tropomyosin and total erythrocyte ghost proteins were included in the blots. Again, we find specific binding of the ¹²⁵I-brain 4.1 to spectrin. One characteristic of erythrocyte 4.1 is that it binds to both α - and β -subunits of erythrocyte spectrin⁷. Blots in which ¹²⁵I-brain 4.1 was incubated with isolated α - and β -subunits of brain spectrin showed specific binding to both. Preliminary results indicate that brain 4.1 and erythrocyte 4.1 compete for the same site on brain spectrin, as brain 4.1 displaces binding of ¹²⁵I-brain spectrin to erythrocyte 4.1 on blots (not shown). The interaction between brain 4.1 and brain spectrin is also demonstrable by affinity chromatography (not shown).



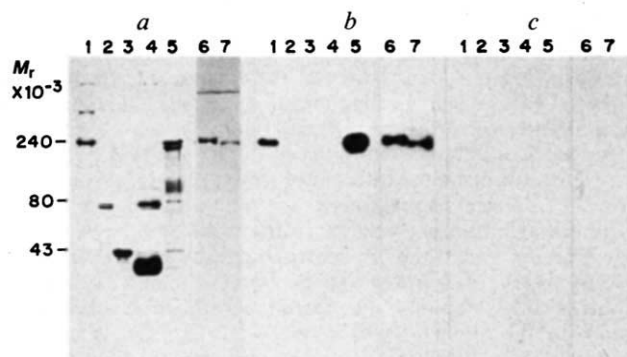


Fig. 3 ^{125}I -brain 4.1 specifically binds to erythrocyte and brain spectrin, α - and β -subunits. After SDS-gel electrophoresis on 5–15% gels, proteins were either stained with Coomassie blue (a) or electrophoretically transferred to nitrocellulose (b, c). Lanes 1, pig brain spectrin; 2, pig brain 4.1; 3, rabbit skeletal muscle actin; 4, rabbit skeletal muscle tropomyosin; 5, human erythrocyte ghosts; 6, brain spectrin α -subunit; 7, brain spectrin β -subunit. Actin, tropomyosin and ghost proteins were included as additional controls for nonspecific interaction between acidic proteins and a basic protein (brain 4.1). Brain spectrin and its subunits were prepared as described previously²⁷. Nitrocellulose replicas were incubated with ^{125}I -brain 4.2 (10 nM, 2×10^6 c.p.m. ml^{-1}) in blot buffer (4% (w/v) bovine serum albumin (Pentex), 0.15 M NaCl, 10 mM Naphosphate, 1 mM EDTA, 0.2% Triton X-100, 1 mM NaN_3 , pH 7.1) for 18 h at 4°C. In c, the ^{125}I -brain 4.1 had been denatured by heating to 60°C for 20 min before incubation with the nitrocellulose paper. Unbound ^{125}I -brain 4.1 was removed by washing the replicas five times with blot buffer lacking albumin.

Brain and erythrocyte 4.1 are the products of different genes, with distinct two-dimensional peptide maps (not shown). Similarly, mammalian brain spectrin^{20,21} and brain ankyrin²² have fingerprints distinct from their erythrocyte counterparts. However, the two components of brain 4.1 give rise to nearly identical fingerprints (not shown). Limited chymotryptic digests of brain 4.1 resolved on SDS gels show ~10 peptides which cross-react with erythrocyte 4.1, suggesting a substantial degree of antigenic homology (not shown).

Like erythrocyte 4.1 (ref. 23), brain 4.1 is a substrate for cAMP-dependent kinase. When brain membranes were incubated with [γ - ^{32}P]ATP and 10 μM cAMP, a pair of polypeptides became labelled which are specifically immunoprecipitable with anti-brain 4.1 antibody. In the absence of cAMP, these polypeptides do not become appreciably labelled (data not shown).

We conclude that the protein we have purified is closely related to erythrocyte 4.1, because: (1) it has cross-reactivity with anti-4.1 IgG; (2) it consists of two basic, sequence-related polypeptides; (3) it binds spectrin; and (4) it is a substrate for cAMP-dependent protein kinases.

The properties of brain 4.1 are strikingly similar to those of synapsin I, a protein associated with synaptic vesicles²⁴. Synapsin I was initially identified as a major substrate in brain for cAMP-dependent protein kinase¹⁵ and is a doublet of $M_r \sim 80,000$ on SDS gels²⁵.

The physical properties of brain 4.1 and synapsin I are similar: synapsin I is basic (pI 10.3) with two polypeptide components that give rise to similar peptide maps²⁵. Both brain 4.1 and synapsin I²⁶ are released from brain membranes in solutions of >0.15 M KCl. Immunoprecipitated ^{32}P -brain 4.1 co-migrates on SDS gels with a pair of ^{32}P -polypeptide that seem to be those defined by Johnson *et al.*¹⁵ as synapsin I (not shown).

Standard synapsin I, prepared by extraction of brain membranes of pH 3, isoelectric precipitation of contaminants at pH 6, followed by hydroxylapatite and gel-filtration chromatography²⁵ (a gift of Dr H. C. Palfrey) has identical electrophoretic mobility on SDS gels to brain 4.1 (Fig. 4). Synapsin I and brain 4.1 were immunologically cross-reactive, as they both reacted on SDS gel blots with anti-red cell 4.1 and anti-brain 4.1 IgG, but not

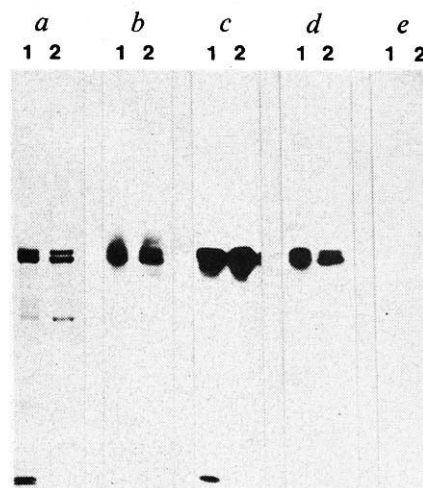


Fig. 4 Synapsin I cross-reacts with anti-RBC 4.1 and anti-brain 4.1 antibodies and binds ^{125}I -spectrin. Samples of bovine synapsin I (lanes 1) and bovine brain 4.1 (lanes 2) were electrophoresed on polyacrylamide gels which were stained with Coomassie blue (a) or transferred to nitrocellulose paper (b–e). The nitrocellulose paper was reacted²⁷ with: b, anti-RBC 4.1 and ^{125}I -protein A; c, anti-brain 4.1 and ^{125}I -protein A; d, ^{125}I -brain spectrin (20 nM, 2×10^6 c.p.m. ml^{-1}); and e, ^{125}I -brain spectrin in the presence of a 40-fold excess of unlabelled spectrin.

with non-immune IgG. Furthermore, they both bind ^{125}I -brain spectrin on gel blots, the binding being displaceable with unlabelled spectrin.

Finally, two-dimensional ^{125}I -iodo peptide maps of synapsin I and brain 4.1 are identical in all but two (minor) spots out of some 30 visible on the maps (Fig. 5), the variant spots differing only in their mobilities on the plates. The variations in the two spots need not imply variations in primary structure, but may reflect different post-translational modifications. Differing phosphorylation states, or modification to the proteins that occurred during their preparation, might give rise to such variation. The proteins clearly have very closely related amino-acid sequences and we believe that they are one and the same.

Synapsin I has previously only been studied as a substrate for protein kinases and its function was unknown. Here we report the first evidence that synapsin is a spectrin-binding protein which seems to be related to erythrocyte membrane skeletal protein 4.1.

An estimate of the stoichiometry of spectrin and synapsin I can be made: in whole brain homogenates, spectrin dimer ($M_r = 500,000$) comprises 2.4% of the total protein²⁷ and synapsin I ($M_r = 75,000$) 0.4% (ref. 28). These values correspond to 48 pmol spectrin and 53 pmol synapsin per mg total protein, that is, close to a ratio of 1:1. This stoichiometry is similar to that of the corresponding proteins in the erythrocyte²⁹.

The subcellular organization of spectrin in neurones has not yet been determined precisely. However, both spectrin and synapsin I are enriched in synapses^{24,30} and are prominent components of isolated synaptosomes^{26,31}. Spectrin is also a prominent component of the postsynaptic density³². Initial work suggested that synapsin I was a component of the postsynaptic density³³, but DeCamilli *et al.*²⁴ concluded it was not, although they could not eliminate the possibility that their antibody was excluded from the postsynaptic density. Bearing in mind that synapsin is present stoichiometrically with spectrin, this point is open to re-examination. Furthermore, recent work^{12,34} suggests that although proteins immunoreactive with erythrocyte spectrin and 4.1 (presumably brain spectrin and synapsin) co-localize in neurones, they occur predominantly in the neuronal perikarya.

A spectrin-synapsin linkage may play a part in neurotransmission. Spectrin may afford a linkage (via synapsin) between

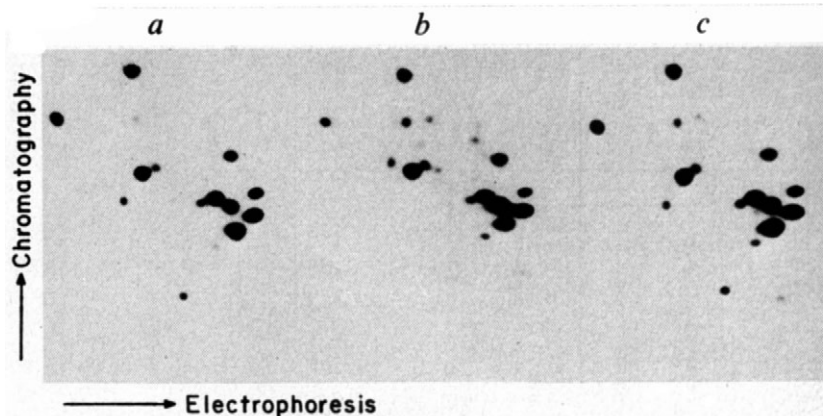


Fig. 5 Brain 4.1 and synapsin I have very similar peptide maps. *a*, Brain 4.1; *b*, synapsin; *c*, a mixture of brain 4.1 and synapsin. Peptide maps of bovine brain 4.1 and bovine synapsin I were prepared by the method described in ref. 27.

synaptic vesicles and the plasma membrane. However, spectrin has also been identified as an axonally transported protein, sometimes associated with membrane-bounded organelles³⁵. Thus, an additional or alternative role for spectrin and synapsin might encompass linkage (or regulation of linkage) of synaptic vesicles to cytoskeletal elements. Such cytoskeletal elements need not be only actin. Although erythrocyte 4.1 seems to enhance the binding of brain spectrin to actin³⁶, our preliminary results have not so far revealed such a role for synapsin. However, synapsin may bind microtubules (unpublished data), which are also present in synaptosomes³¹.

Erythrocyte 4.1 is able to bind at least one transmembrane protein. Binding of synapsin to a transmembrane protein of the synaptic vesicle may be of importance in regulation of neurotransmitter secretion, as synapsin is released from its membrane binding site on phosphorylation of synapsin by cAMP- or Ca^{2+} /calmodulin-dependent protein kinases²⁶, which phosphorylate synapsin in response to hormonal or electrical stimulation of nerves³⁷.

This work was supported by Damon Runyon-Walter Winchell Cancer Fund Fellowship DRG700 and by NIH grants RO1 AM29808, RO1 GM33996 and a RCDA KO4 AM00926. We thank Arlene Daniel for preparation of the manuscript.

Received 2 October 1984; accepted 11 March 1985.

1. Baines, A. J. *Nature* **301**, 377-378 (1983).
2. Baines, A. J. *Nature* **312**, 310-311 (1984).
3. Bennett, V. A. *Rev. Biochem.* (in the press).
4. Ungewickell, E., Bennett, P. M., Calvert, R., Ohanian, V. & Gratzner, W. B. *Nature* **280**, 811-814 (1979).
5. Tyler, J., Reinhart, B. & Branton, D. *J. biol. Chem.* **255**, 7034-7039 (1980).
6. Ohanian, V. *et al. Biochemistry* **23**, 4416-4420 (1984).
7. Cohen, C. M. & Langley, R. C. *Biochemistry* **23**, 4488-4495 (1984).
8. Anderson, R. A. & Lovrien, R. E. *Nature* **307**, 655-658 (1984).
9. Mueller, T. *Fedn Proc.* **43**, 1849a (1984).
10. Pasternak, G., Anderson, R. A. & Marchesi, V. T. *J. biol. Chem.* **260**, 3676-3683 (1985).
11. Cohen, C. M., Foley, S. F. & Korsgren, C. *Nature* **299**, 648-650 (1982).
12. Spiegel, J. E., Beardsley, D. S., Southwick, F. S. & Lux, S. E. *J. Cell Biol.* **99**, 886-893 (1984).
13. Goodman, S. R., Casoria, L. A., Coleman, D. B. & Zagon, I. G. *Science* **224**, 1433-1436 (1984).
14. Granger, B. L. & Lazarides, E. *Cell* **37**, 595-607 (1984).
15. Johnson, E. M., Ueda, T., Maeno, H. & Greengard, P. *J. biol. Chem.* **257**, 5650-5652 (1972).
16. Sieghart, W., Forn, J. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 2475-2479 (1979).
17. Laemmli, U. K. *Nature* **227**, 680-685 (1970).
18. Gorbunoff, M. J. *Analyt. Biochem.* **136**, 425-432 (1984).
19. Fairbanks, G., Steck, T. L. & Wallach, D. F. H. *Biochemistry* **10**, 2606-2617 (1970).
20. Bennett, V., Davis, J. & Fowler, W. E. *Nature* **299**, 126-131 (1982).
21. Glenney, J. R. & Glenney, P. *Eur. J. Biochem.* **144**, 529-539 (1984).
22. Davis, J. & Bennett, V. *J. biol. Chem.* **259**, 1874-1881 (1984).
23. Plut, D. A., Hosey, M. M. & Tao, M. *Eur. J. Biochem.* **82**, 333-337 (1978).
24. DeCamilli, P., Harris, S. M., Huttner, W. B. & Greengard, P. *J. Cell Biol.* **96**, 1355-1373 (1983).
25. Ueda, T. & Greengard, P. *J. biol. Chem.* **252**, 5155-5163 (1977).
26. Huttner, W. B., Schiebler, W., Greengard, P. & DeCamilli, P. *J. Cell Biol.* **96**, 1374-1388 (1983).
27. Davis, J. & Bennett, V. *J. biol. Chem.* **258**, 7757-7766 (1983).
28. Goetz, S. E., Nestler, E. J., Chehrizi, B. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **78**, 2130-2134 (1981).
29. Pinder, J. C. & Gratzner, W. B. *J. Cell Biol.* **96**, 768-775 (1983).
30. Koenig, E. & Repasky, E. *J. Neurosci.* **5**, 705-714 (1985).
31. Hesketh, J. E., Thierse, D. & Aunis, D. *Biochem. biophys. Res. Commun.* **115**, 437-443 (1983).
32. Carlin, R. K., Bartelt, D. C. & Siekevitz, P. *J. Cell Biol.* **96**, 443-448 (1983).
33. Bloom, F. E. T., Ueda, N. T., Battenberg, E. & Greengard, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 5982-5986 (1979).
34. Zagon, I. S., McLaughlin, P. J. & Goodman, S. R. *J. Neurosci.* **4**, 3089-3100 (1984).
35. Koenig, E., Kinsman, S., Repasky, E. & Sultz, L. *J. Neurosci.* **5**, 715-729 (1985).
36. Burns, N. R., Ohanian, V. & Gratzner, W. B. *FEBS Lett.* **153**, 165-168 (1983).
37. Nestler, E. J. & Greengard, P. *Nature* **296**, 452-454 (1982).
38. Olmstead, J. B. *J. biol. Chem.* **256**, 11955-11957 (1981).

Expression of human growth hormone-releasing factor in transgenic mice results in increased somatic growth

Robert E. Hammer*, Ralph L. Brinster*, Michael G. Rosenfeld†, Ronald M. Evans‡ & Kelly E. Mayo‡

* Laboratory of Reproductive Physiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

† School of Medicine, University of California, San Diego, California 92023, USA

‡ Molecular Biology and Virology Laboratory, The Salk Institute for Biological Studies, PO Box 85800, San Diego, California 92138, USA

The neurohumoral regulation of growth hormone secretion is mediated in part by two hypothalamic peptides that reach the anterior pituitary via the hypothalamo-hypophyseal portal blood system¹. Somatostatin inhibits the release of growth hormone², whereas growth hormone-releasing factor (GRF) positively regulates both growth hormone synthesis^{3,4} and secretion^{5,6}. Two forms of human GRF, 40 and 44 amino acids long, have been characterized from extra-hypothalamic tumours^{7,8} as well as from the hypothalamus⁹. Analysis of human GRF complementary DNA^{10,11} and genomic¹² clones indicates that the GRF peptides are first synthesized as a 107- or 108-amino-acid precursor protein. To examine the physiological consequences of GRF expression, we have established strains of transgenic mice containing a fusion gene including the promoter/regulatory region of the mouse metallothionein-I (MT-I) gene¹³ and the coding region of the human GRF gene¹². We report that expression of the human GRF precursor protein in these animals results in measurable levels of human GRF and increased levels of mouse growth hormone in plasma and accelerated growth rates relative to control littermates. These results demonstrate a direct role for GRF in the positive regulation of somatic growth. Unexpectedly, female transgenic mice carrying the MT-GRF fusion gene are fertile, in contrast to female transgenic mice expressing human or rat growth hormone, which are generally infertile. These transgenic mouse strains should provide useful animal models for the study of several types of human growth disorders.

It was shown previously that the hormonal cascade controlling somatic growth could be manipulated by the germline introduction of foreign genes. Expression of either rat¹⁴ or human¹⁵ growth hormone (GH) in transgenic mice results in increased growth rates, giving rise to large animals. Because GRF is believed to be an important physiological regulator of growth hormone in animals, we sought to analyse further its role in growth control by creating strains of animals that express supraphysiological levels of GRF. Neither rat nor human growth hormone genes are themselves expressed in transgenic animals^{16,17}; this deficiency has been overcome by fusing these structural genes to an expressed heterologous gene promoter.

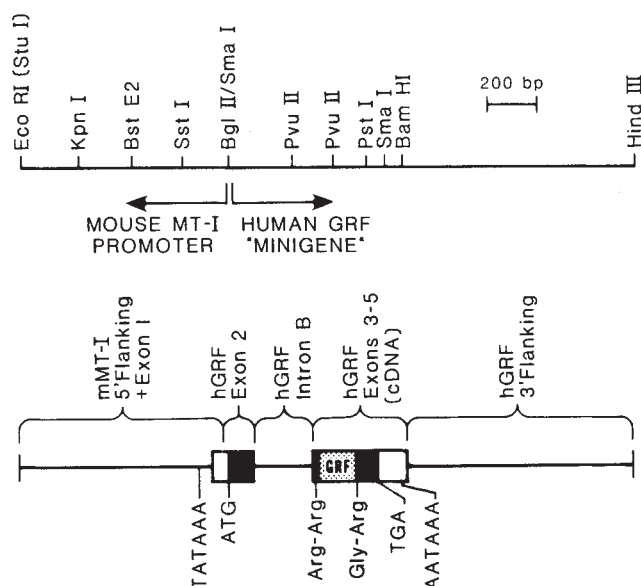


Fig. 1 Structure of the mouse metallothionein-I/human growth hormone-releasing factor fusion gene. The upper portion of the figure is a restriction map for common 6-bp cutters. *BglII/SmaI* indicates the site of the fusion between the mouse MT-I promoter and the human GRF minigene. The lower portion of the figure is a schematic representation of the 2.5-kb *EcoRI-HindIII* fragment of DNA used for microinjection. Open boxes represent 5' and 3' untranslated exon regions, and shaded boxes represent the coding exon regions. The derivations of various parts of the fusion gene are indicated. The fusion occurs in the 5' untranslated regions of both genes. The large exon labelled 3-5 was constructed by replacing part of the human GRF gene with a human GRF cDNA, effectively eliminating intron C (2.4 kb) and intron D (3.0 kb) of the human GRF gene¹². Consensus sequences involved in transcription initiation, translation initiation and termination, polyadenylation, and proteolytic processing of the GRF peptide from its precursor are shown.

The mouse metallothionein-I gene promoter offers several advantages for this type of approach; it is expressed at fairly high levels in most mouse tissues¹⁸, it can be regulated by heavy metals¹⁸, and both expression and regulation are retained in gene transfer experiments¹⁹⁻²¹. Because of the possibility that the human GRF gene, like the human growth hormone gene, might not be expressed in transgenic animals, we initially chose to use the heterologous mouse MT-I promoter to express human GRF.

The structure of the mouse MT-I/human GRF fusion gene construct used (referred to as MT-GRF) is shown in Fig. 1. A 770-base pair (bp) fragment of the mouse MT-I gene, including sequences responsible for metal-inducibility and transcription initiation, was fused to a human GRF minigene which includes the entire coding region of the GRF precursor protein. The human GRF minigene was created by combining cDNA and genomic clones such that the 10-kilobase (kb) human GRF gene, which normally includes five exons¹², has been reduced to <1 kb and retains a single intron of 230 bp (see Fig. 1). Analysis of human GRF cDNA clones suggests that the GRF precursor protein can consist of either 107 or 108 amino acids, differing in the presence or absence of serine 103 (ref. 11). DNA sequence analysis of the human GRF gene¹² indicates that this difference may be explained by alternative RNA processing. The cDNA used to construct the MT-GRF fusion gene should encode only the 108-amino-acid form of the GRF precursor protein¹⁰. To determine whether the MT-GRF fusion gene could be expressed and regulated correctly, it was introduced into cultured mouse fibroblast cells using CaPO_4 -mediated DNA transfection²² and co-selection for neomycin resistance conferred by the vector pSV2-neo²³. Several stable cell lines that were generated in this manner express a MT-GRF fusion mRNA of the expected size whose abundance is increased by metal treatment, and accumu-

Table 1 Expression of the MT-GRF fusion gene in transgenic mice

Mouse	No. of gene copies per cell	Liver GRF mRNA (relative amount)	Plasma hGRF (ng ml ⁻¹)	Plasma mGH (ng ml ⁻¹)	Growth (ratio)
♂760-5	<1	3	24	1,051	1.35
♂762-3	2	0	<10	21	0.93
♀762-5	4	63	26	141	1.33
♀765-2	2	102	207	415	1.24
♂800-1	1	0	<10	149	1.12
♀800-8	1	3	14	402	1.05
♂801-3	3	0	<10	38	0.97
♀801-5	8	38	99	809	1.35
♂801-9	1	4	20	541	1.42
♀802-3	5	1	45	913	1.51
♀803-4	10	118	263	1,095	1.41
♀803-5	1	5	24	166	1.45
♂803-6	10	16	50	302	1.24
♀803-7	20	*	*	*	1.36

A 2.5-kb *EcoRI-HindIII* fragment containing the MT-GRF fusion gene was microinjected into the male pronucleus of fertilized F₂ hybrid mouse eggs and the eggs implanted into the oviducts of pseudopregnant recipients²⁴. At weaning, DNA was isolated from a piece of tail and analysed for the fusion gene by dot-blotting using a human GRF cDNA as probe¹⁰. Positive animals were re-analysed by Southern DNA blotting and copy number determined by comparison with a standard curve generated from known amounts of MT-GRF plasmid DNA. Positive animals were maintained on water containing 25 mM ZnSO_4 . RNA was isolated following partial hepatectomy and analysed by Northern blotting using a human GRF cDNA probe. Autoradiograms were densitometrically scanned to determine relative RNA amounts. Plasma GRF and GH were determined by radioimmunoassay of serum samples from animals at 9 weeks old. The GH radioimmunoassay was able to detect 16 ng ml⁻¹ serum GH (two control animals had 16 and 46 ng ml⁻¹ GH) and the GRF radioimmunoassay was able to detect 10 ng ml⁻¹ serum GRF (two control animals had <10 ng ml⁻¹ GRF). The growth ratio was determined at 9 weeks of age and represents a comparison with age- and sex-matched littermates.

* ♀803-7 died prematurely. Although we were able to measure liver GRF mRNA, the amount could not be quantitated because of RNA degradation in the postmortem liver. Plasma GRF and GH levels could not be determined accurately for the same reason.

late radioassayable human GRF (hGRF) in the culture medium (K.E.M., unpublished results).

The MT-GRF fusion gene (~1,600 molecules) was microinjected into the male pronucleus of 350 F₂ hybrid eggs (obtained by mating C57BL/6 × SJL hybrid adults) and the eggs were transferred into the oviducts of pseudopregnant recipients as described elsewhere²⁴. Fifty-nine animals developed from the microinjected eggs; at the time of weaning, DNA was isolated from a piece of tail taken from these animals. The presence of the foreign MT-GRF gene was detected by hybridization to a human GRF cDNA clone¹⁰ using conditions in which cross-hybridization with the mouse GRF gene is negligible. Fourteen animals carried the MT-GRF fusion gene; these animals were maintained on water with 25 mM ZnSO_4 to enhance expression of the fusion gene¹⁵. Table 1 summarizes the expression of the MT-GRF fusion gene in these 14 transgenic mice; 11 of the 14 animals express the MT-GRF messenger RNA in the liver, a major site of metallothionein expression¹⁸. The amount of the fusion mRNA in the livers of these mice varies by as much as 100-fold, and does not seem to be strongly correlated with the number of copies of the fusion gene in each animal. Animals expressing the fusion gene have measurable levels of human GRF in their serum, and have increased levels of serum growth hormone (Table 1). Ten of the mice showed significant increases in growth at 9 weeks old, and were 25-50% larger than control littermates. At maturity, some of the MT-GRF animals were nearly twice the size of control animals. Examination of the data in Table 1 reveals that there is no strong correlation between GRF levels, GH levels and growth.

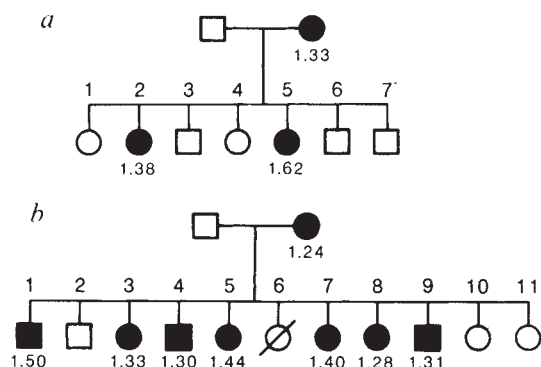


Fig. 2 Pedigree analysis of MT-GRF transgenic mice. Mice that carry the MT-GRF fusion gene (determined by DNA blotting of tail snips) are indicated by the solid symbols. Squares represent males and circles represent females. The animal number is given above each symbol; the numbers below the solid symbols indicate weight ratios at 9 weeks old compared with age- and sex-matched littermates. *a*, Pedigree of MT-GRF mouse 762-5; *b*, that of MT-GRF mouse 765-2. Mouse 765-2-6 died but was previously scored as a non-carrier.

To determine whether the MT-GRF fusion gene and the large size are heritable, two of the MT-GRF founder animals were bred with control animals; offspring from these matings were analysed for the presence of the MT-GRF fusion gene and for their rate of growth. Figure 2 shows two pedigrees demonstrating that both the fusion gene and the 'large' phenotype are inherited by ~50% of the total offspring. DNA analysis showed that all positive offspring carried the same number of copies of the fusion gene as the corresponding parent (data not shown). Four other females were also mated and produced litters. Thus, all six of the females tested were fertile (two were killed before breeding). In addition, male MT-GRF mice 760-5 and 801-9 (see Table 1) sired litters that included offspring carrying the fusion gene and displaying the large phenotype, indicating that both males and females carrying the MT-GRF fusion gene are fertile and transmit the gene. We now know that both the MT-GRF fusion gene and the large phenotype are inherited by ~50% of the offspring examined in the F_2 generation of animals.

To determine whether expression of the MT-GRF fusion gene in transgenic mice demonstrates a tissue specificity similar to that of the endogenous MT-I gene, we analysed both MT-I and MT-GRF RNAs in six tissues from either a control mouse or an MT-GRF mouse (female 765-2-3) (see F_1 in Fig. 2b); the

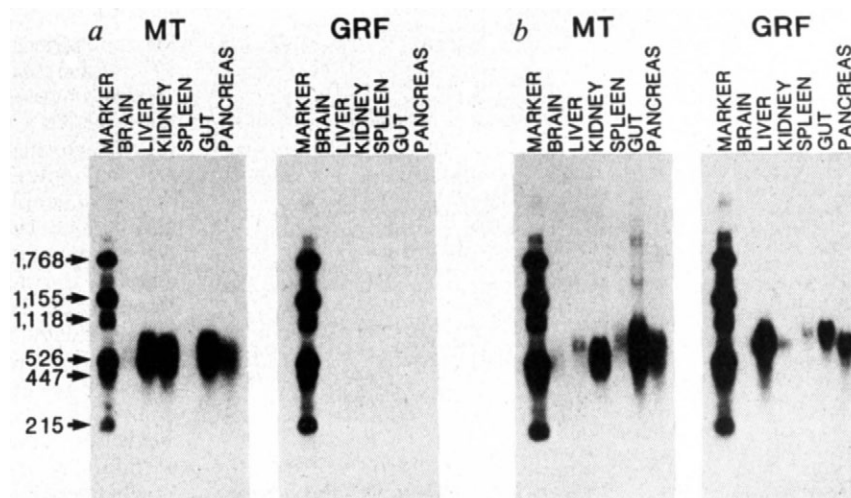
results are presented in Fig. 3. The control animal expressed MT-I mRNA at high levels in liver, kidney, gut and pancreas, and at lower levels in brain and spleen. This agrees well with the known tissue distribution of MT-I mRNA¹⁸. The control animal did not express detectable hGRF mRNA in any tissue. Although the hGRF probe might detect endogenous mouse GRF mRNA in the brain, the low abundance of GRF mRNA even in the hypothalamus (<0.01%) makes detection unlikely. The transgenic mouse 765-2-3 showed a tissue distribution of MT-I mRNA much like that of the control mouse with one exception; the amount of liver MT-I mRNA was substantially reduced (Fig. 3). The reason for this is unclear, but it cannot be explained by RNA degradation and is a reproducible result in this animal. In the transgenic mouse, the MT-GRF fusion mRNA is expressed at high levels in the liver, gut and pancreas, at low levels in kidney and spleen, and was not detected in the brain (Fig. 3). Although the expression of the fusion gene in this animal demonstrates a tissue specificity similar to that of the mouse MT-I gene, there are clear and reproducible differences, most noticeably the low level of expression of MT-GRF mRNA in the kidney. We have recently examined the pattern of MT-GRF gene expression in a second pedigree (801-5-9) and have found that although the general pattern of tissue specificity parallels that of mouse MT-I, there are subtle differences particular to this animal and distinct from the MT-GRF animal 765-2-3 (data not shown). Similar results were obtained on examination of the tissue specificity of metallothionein/growth hormone fusion genes¹⁵. In this case, an extensive survey of multiple animals showed that expression varied among different tissues and different animals, suggesting that factors such as the site of integration probably influence expression of the foreign gene.

Although many tissues apparently express the MT-GRF mRNA and presumably synthesize the GRF precursor protein, we do not know which tissues are capable of proteolytically processing this precursor to generate the mature GRF peptide. The antibody used to detect plasma GRF would detect both the precursor and the mature peptide¹⁰. It seems likely that such processing must occur to some extent, as the free N-terminus of the GRF peptide is known to be required for biological activity^{7,8}. We are presently using chromatographic techniques to determine what percentage of the immunoassayable GRF in the plasma is mature peptide as opposed to precursor.

Our results demonstrate that expression of supraphysiological levels of GRF in mice leads to increased somatic growth. Conversely, it has been shown that passive immunization of rats with antibodies against GRF can inhibit somatic growth²⁵. Taken

Fig. 3 Expression of human GRF mRNA in MT-GRF transgenic mice. *a*, Results from a control mouse; *b*, results from MT-GRF transgenic mouse 765-2-3. Each panel shows Northern blots of RNAs isolated from six different tissues. Duplicate blots from each animal were probed with either mouse MT-I (mMT-I) or hGRF probes as indicated. Size markers (in nucleotides) are denatured *Hind*III-cut simian virus 40 DNA.

Methods. Both animals were maintained on water with 76 mM $ZnSO_4$ for several weeks before killing. The indicated organs were removed and frozen at $-70^\circ C$ until use. Total RNA was prepared by homogenization in guanidine isothiocyanate and centrifugation through caesium chloride as described²⁹; 5 μg of each RNA was denatured and electrophoresed on formaldehyde/1.5% agarose gels³⁰ and the RNA transferred to nitrocellulose filters as described³¹. Filters were probed with either a mouse MT-I genomic clone, pSH¹³, or a human GRF cDNA clone, phGRF-54 (ref. 10). Probes were labelled by nick-translation using [α -³²P]dCTP (NEN). Exposure time of the autoradiograms ~16 h.



together, these data suggest strongly that GRF has an important role in the promotion of growth. Transgenic mice carrying the MT-GRF fusion gene have elevated levels of plasma growth hormone and grow at a rate 25–50% greater than that of control littermates; this is significantly less than the increased growth observed in transgenic mice expressing growth hormone, many of which grow to twice the normal size^{14,15}. However, in the transgenic animals expressing growth hormone, most body tissues express the gene and have the potential to make growth hormone, whereas in our experiments growth hormone is only made in the anterior pituitary and thus may be rate-limiting. An unexpected finding was that females expressing the MT-GRF fusion gene are fertile, although transgenic female mice expressing the growth hormone gene are generally infertile¹⁶; this suggests that the enhancement of growth in the MT-GRF females is more physiological, perhaps because the effect is mediated through endogenous somatotropes.

We examined the pituitaries of two MT-GRF transgenic animals that express the fusion gene at high levels (♀762-5 and ♀765-2-3). In both cases the pituitary was at least three times normal size and showed signs of marked hyperplasia. The condition of these animals is remarkably similar to the clinical findings in patients with certain types of human growth disorders characterized by elevated levels of growth hormone in serum and pituitary hyperplasia. Rather than pituitary adenoma, the more common cause of such symptoms, a tumour is found at a remote site, often a pancreatic or bronchial carcinoma^{26–28}. These tumours do not produce growth hormone, yet on resection plasma GH levels fall. Such rare tumours are now recognized to be ectopic sources of GRF^{7,8} and the MT-GRF transgenic mice seem to provide a good animal model for this class of growth disorders. The finding of pituitary hyperplasia suggests that GRF, in addition to regulating GH production, might have a role in somatotroph proliferation.

We thank Richard Palmiter for providing the mouse MT-I promoter and for continued interest in the project, Gail Cerelli, Paul Fallon and Myrna Trumbauer for technical assistance, the laboratory of Wylie Vale for GRF and GH radioimmunoassays, and Marijke terHorst for preparing the manuscript. This work was supported by grants from the NIH and NIAMDD to R.M.E., and from the NIH to R.L.B. R.E.H. was supported by NIH training grant HD07155 and K.E.M. by Damon Runyon-Walter Winchell Cancer Fund Fellowship DRG-648.

Received 14 January; accepted 26 March 1985.

- Martin, J. B. *Physiologist* **22**, 23–29 (1979).
- Brazeau, P. et al. *Science* **179**, 77–79 (1973).
- Baringa, M. et al. *Nature* **306**, 84–85 (1983).
- Glick, G. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1553–1555 (1984).
- Brazeau, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7909–7913 (1982).
- Vale, W., Vaughan, J., Yamamoto, G., Spiess, J. & Rivier, J. *Endocrinology* **112**, 1553–1555 (1983).
- Rivier, J., Spiess, J., Thorner, M. & Vale, W. *Nature* **300**, 276–278 (1982).
- Guillemin, R. et al. *Science* **218**, 585–587 (1982).
- Ling, N. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4302–4306 (1984).
- Mayo, K. E., Vale, K. E., Rivier, J., Rosenfeld, M. G. & Evans, R. M. *Nature* **306**, 86–88 (1983).
- Gubler, U. et al. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4311–4314 (1983).
- Mayo, K. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 63–67 (1985).
- Durnam, D. M., Perrin, F., Gannon, F. & Palmiter, R. D. *Proc. natn. Acad. Sci. U.S.A.* **77**, 6511–6515 (1980).
- Palmiter, R. D. et al. *Nature* **300**, 611–615 (1982).
- Palmiter, R. D., Norstedt, G., Gellinas, R. E., Hammer, R. E. & Brinster, R. L. *Science* **222**, 809–814 (1983).
- Hammer, R. E., Palmiter, R. D. & Brinster, R. L. *Nature* **311**, 65–67 (1984).
- Wagner, E. F., Covarrubias, L., Stewart, T. A. & Mintz, B. *Cell* **35**, 647–655 (1983).
- Durnam, D. M. & Palmiter, R. D. *J. biol. Chem.* **256**, 5712–5716 (1981).
- Mayo, K. E., Warren, R. & Palmiter, R. D. *Cell* **29**, 99–108 (1982).
- Brinster, R. L., Chen, H. Y., Warren, R., Santhya, A. & Palmiter, R. D. *Nature* **296**, 39–42 (1982).
- Palmiter, R. D., Chen, H. Y. & Brinster, R. L. *Cell* **29**, 701–710 (1982).
- Graham, F. L. & Van der Eb, A. J. *Virology* **52**, 456–467 (1973).
- Southern, P. J. & Berg, P. *J. molec. appl. Genet.* **1**, 327–341 (1982).
- Brinster, R. L. et al. *Cell* **27**, 223–231 (1981).
- Wehrenberg, W. B., Bloch, B. & Phillips, B. *J. Endocrinology* **115**, 1218–1220 (1984).
- Frohman, L. A. & Szabo, M. in *Physiopathology of Endocrine Diseases and Mechanisms of Hormone Action*, 259–271 (Liss, New York, 1981).
- Thorner, M. O. et al. *J. clin. Invest.* **70**, 965–977 (1982).
- Asa, S. L. et al. *J. clin. Endocr. Metab.* **58**, 796–803 (1984).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294–5299 (1979).
- Meinkoth, J. & Wahl, G. M. *Analyt. Biochem.* **138**, 267–284 (1984).
- Thomas, P. *Proc. natn. Acad. Sci. U.S.A.* **77**, 5201–5205 (1980).

Phagosome acidification blocked by intracellular *Toxoplasma gondii*

L. David Sibley*, Earl Weidner* & James L. Krahenbuhl†

* Department of Zoology and Physiology, Louisiana State University, Baton Rouge, Louisiana 70803, USA

† Department of Immunology Research, National Hansen's Disease Center, Carville, Louisiana 70721, USA

Toxoplasma gondii belongs to a group of highly virulent intracellular parasites that reside in host cell vacuoles which resist typical phagosome-lysosome fusion¹. Live *Toxoplasma* replicate prodigiously within modified phagocytic vacuoles formed during invagination of the host plasma membrane^{2,3}. In contrast, heat-killed *Toxoplasma* or specific antibody (heat-inactivated)-coated live *Toxoplasma*-containing vacuoles readily undergo lysosome fusion and digestion in normal macrophages^{2,4}. Of newly recognized significance to *Toxoplasma* survival is the microbicidal effect of phagosome acidification, which reportedly can occur independently of fusion with other acidic vesicles^{5–8}. We report here that modified live *Toxoplasma*-containing vacuoles fail to acidify in normal macrophages, as indicated by the sensitive pH probe fluorescein. In contrast, when live *Toxoplasma* are coated with specific antibody (heat-inactivated), they trigger phagosome acidification when entering normal macrophages. A similar acidification is observed when normal phagocytes ingest dead *Toxoplasma*. Extracellular *Toxoplasma* are highly susceptible to acidic pH conditions, indicating that the acidification block in the modified vacuoles may be important for intracellular survival.

The endogenous acidification of macrophage phagosomes may involve two independent mechanisms that operate before acidic vesicle fusion. First, NADPH-oxidase activity, associated with the phagocytic respiratory burst, forms superoxide radicals and delivers H⁺ ions to the phagosome⁵. Second, phagosome acidification may involve the Mg-ATPase H⁺-ion pump identified in pinosomes⁹ and ligand-receptor endosomes^{10–15} also formed by plasma membrane invagination. Given the potential for similar acidic changes in newly formed phagosomes⁶ and the pH susceptibility of *Toxoplasma*^{16,17}, we investigated how the *in situ* pH of macrophage vacuoles containing *Toxoplasma* affected intracellular survival.

The fate of live compared with heat-killed or antibody-coated *Toxoplasma* in normal or activated mouse peritoneal macrophages was evaluated microscopically at intervals after infection *in vitro*. Live *Toxoplasma* here refers to extracellular parasites freshly collected from mouse ascites¹⁸, representing the maximum viability obtainable. Acridine orange was used to prelabel acidic cell compartments, including lysosomes of macrophages cultured on coverslips¹⁹. One hour after infection, 80% of live *Toxoplasma*-containing vacuoles blocked fusion of acridine orange-labelled vesicles in normal macrophages. However, 86% of heat-killed (60 °C for 10 min) or antibody (heat-inactivated)-coated live *Toxoplasma* vacuoles fused with acridine orange-labelled vesicles within 1 h of phagocytosis by normal macrophages. In activated macrophages from immune mice⁹ which were cultured *in vitro* for 8 h, 85% of live *Toxoplasma* vacuoles fused with acridine orange compartments, were digested, and thus resemble heat-killed or antibody-coated *Toxoplasma* vacuoles.

To evaluate the cidal effect of acidic pH, extracellular *Toxoplasma* were incubated in media at fixed pH values and then evaluated for their ability to survive in host cells. The viability of live *Toxoplasma* was reduced to ≤60% of control after incubation in media at pH 6.0, as evaluated by plaque formation on fibroblast monolayers²⁰ and fusion with acridine orange-labelled vesicles in normal macrophages (Fig. 1).

Having established the high susceptibility of extracellular *Toxoplasma* to acidic pH, we examined the *in situ* pH of modified

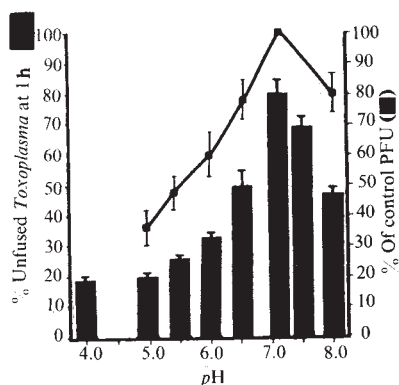


Fig. 1 Viability of *Toxoplasma* as evaluated by ability to block fusion of acidic vesicles (labelled with acridine orange) in macrophages (bars) and plaque formation on fibroblast monolayers (line) (expressed as plaque-forming units, PFU). Significant reduction in *Toxoplasma* viability was observed by both methods when extracellular *Toxoplasma* were exposed to media of pH 6.0.

Methods. RH strain *Toxoplasma* tachyzoites (asexually replicating stages from acute infections in mice) were collected in Hank's balanced salt solution (HBSS) pH 7.2 (ref. 18). After 10 min incubation in isotonic HBSS adjusted to pH values from 4.0 to 8.0 (Orion Research millivolt pH meter 611), *Toxoplasma* were returned to HBSS at pH 7.2 and used for viability assays within 30 min. Normal mouse peritoneal macrophages were cultured in endotoxin-free (≤ 0.01 ng ml $^{-1}$, Limulus Amebocyte Assay, Associates of Cap Cod Inc.) RPMI-1640 with 10% heat-inactivated fetal calf serum (CRPMI, Gibco) on 12-mm-diameter glass coverslips in 24-well plates for 18 h before use. The extent of fusion in normal macrophages labelled with acridine orange¹⁹ was evaluated 1 h after *Toxoplasma* infection ($n = 3$)²⁵. Plaque formation on human fibroblast monolayers²⁰ was evaluated after 7 days and expressed as mean % control (pH 7.2) from triplicate 25-cm 2 flasks.

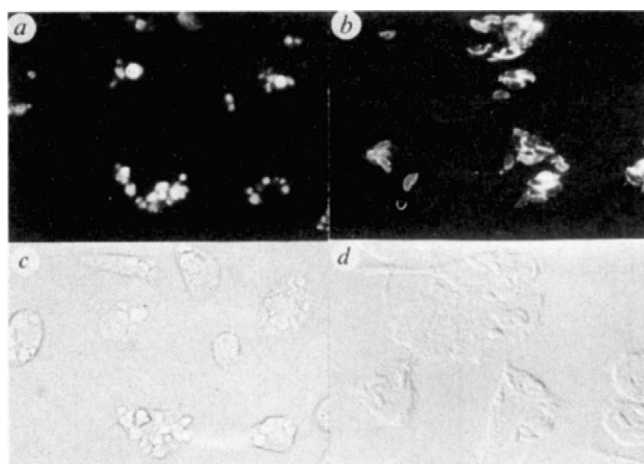


Fig. 2 Macrophage monolayers containing fluorescein-labelled *Toxoplasma*-containing vacuoles photographed on a Leitz Ortholux II epifluorescence microscope using a Nikon HFX 35-mm camera on Kodak Tri-X panchromatic film. *a*, Epifluorescence; *b*, phase contrast; *c*, epifluorescence; *d*, differential interference contrast. Monolayers were pulsed for 5 min with live or heat-killed *Toxoplasma* in CRPMI containing 2 mg ml $^{-1}$ fluorescein dextran (relative molecular mass 20,000; Sigma), rinsed extensively in phosphate-buffered saline (PBS) and recultured. Approximately 80% of phagocytic vacuoles contained visible fluorescein dextran which did not affect block of acridine orange-labelled vesicle fusion or *Toxoplasma* replication in normal macrophages. Spectrofluorometry measurements indicated that pinocytosis of fluorescein dextran accounted for 10% of the total label internalized during *Toxoplasma* infection. *Toxoplasma* labelled with rabbit IgG and FITC-goat anti-rabbit (Sigma) were internalized by macrophage monolayers during 1 min infection challenge.

and heat-killed or antibody-coated *Toxoplasma* vacuoles in macrophage monolayers. *In situ* pH measurements were estimated from a standard curve relating fluorescence intensity (I_0) of intracellular fluorescein-labelled *Toxoplasma* to fixed pH values (s.e. = ± 0.15 unit; see Fig. 3*a* legend). The excitation spectrum of fluorescein is sensitive to pH changes between 4.0 and 8.0 and is independent of fluorochrome concentration when expressed as the I_0 (520 nm) ratio of excitation at 495/450 nm (ref. 21).

Modified live *Toxoplasma*-containing vacuoles were labelled using the fluid phase marker fluorescein dextran present in the media during infection with *Toxoplasma* (Fig. 2*a, b*). Normal macrophage monolayers containing modified vacuoles labelled with fluorescein dextran showed a slight decrease in I_0 during the first hour after infection (Fig. 3*b*). This slight acidification from pH 7.2 to 6.8 is probably due to the 20% of *Toxoplasma* which fuse with acidic cell compartments and the small amount of fluorescein dextran pinocytosed by macrophages during infection (see Fig. 2 legend). Most fluorescein dextran in the modified vacuoles did not undergo decreases in I_0 , demonstrating that live *Toxoplasma* reside in neutral pH vacuoles.

The acidification of heat-killed or antibody-coated *Toxoplasma* vacuoles, which are destined to fuse with lysosomes, was evaluated in macrophages challenged with either heat-killed *Toxoplasma* in the presence of fluorescein dextran (Fig. 2*a, b*) or fluorescein isothiocyanate (FITC)-antibody-coated *Toxoplasma* (Fig. 2*c, d*). Normal macrophage vacuoles containing heat-killed *Toxoplasma* and fluorescein dextran underwent an extensive fall in pH from 7.2 to 5.2 after 1 h (Fig. 3*b*). Consistent with their intracellular fate, a similar rapid acidification of these vacuoles was observed with FITC-antibody-coated live *Toxoplasma* in normal macrophages (Fig. 3*b*). Of considerable interest, live *Toxoplasma* labelled with fluorescein dextran

underwent a rapid acidification in activated macrophages (Fig. 3*b*). We did not detect the previously reported transient alkalization to pH 7.4 at 2–3 min post-phagocytosis, although this may be because we used external media at pH 7.2 rather than 6.8 (ref. 6).

The kinetics of acidification of heat-killed or antibody-coated *Toxoplasma* vacuoles (Fig. 3*b*) indicates that a rapid shift to acidic pH is initiated within the first 3–5 min, starting from pH 7.2 and falling to pH 5.6 at 15 min post-phagocytosis. This rapid acidification rate appears similar to that of other reports on phagosomes of macrophages⁶ and *Amoeba*²². During the initial 15 min post-phagocytosis, the rate of pH change in phagosomes is significantly higher than the rate of fusion with acidic compartments prelabelled with acridine orange (Fig. 4). At 15 min, when acidification is 70% complete, 65% of the vacuoles still show no evidence of acridine orange vesicle fusion (Fig. 4). This pattern is consistent with the model that phagocytic vacuole acidification precedes fusion with other acidic vesicles, including lysosomes⁸.

Several mechanisms may be involved in the rapid acidification of compartments of heat-killed or antibody-coated *Toxoplasma* vacuoles. First, acidification may represent fusion of cytoplasmic vesicles which deliver proton-pumping ability to the phagosome, such as acidosomes described in *Paramecium*²³. However, this explanation would only fit the data in Fig. 4 if the delivery vesicles are not active before vacuole fusion and thus not labelled with acridine orange. In this case, the failure of acidification of live *Toxoplasma* vacuoles represents a block of delivery-vesicle fusion. Alternatively, the observed differences in acidification may result from the degree of endogenous NADPH-oxidase and Mg-ATPase H $^+$ -ion pump activity within the phagosome.

We have demonstrated here that live *Toxoplasma*-containing vacuoles show little capacity for acidification whereas both

antibody-coated and heat-killed *Toxoplasma* are internalized by normal macrophages into phagocytic vacuoles with an inherent ability to acidify. Activated macrophages challenged with live *Toxoplasma* also form phagocytic vacuoles with a marked ability to acidify. Similarly, the entry of live *Toxoplasma* into normal macrophages does not stimulate the oxygen radical production

which is typical of phagocytosis of antibody-coated *Toxoplasma* by normal macrophages and of live *Toxoplasma* phagocytosed by activated macrophages¹⁸. These two observations suggest that differences in plasma membrane enzyme activation during formation of live as opposed to heat-killed or antibody-coated *Toxoplasma* vacuoles are responsible for their ultimate intracellular fate.

Although Fc and C3b receptors may trigger acidification during phagocytosis of antibody-coated *Toxoplasma*, acidification resulting from ingestion of heat-killed *Toxoplasma* by normal macrophages and live *Toxoplasma* by activated macrophages must be initiated by an alternative receptor, as antibody was not present in these preparations. Initiation is either not triggered when live *Toxoplasma* enter normal macrophages, or is inhibited, as demonstrated by the failure of modified live *Toxoplasma*-containing phagosomes to acidify. A similar pattern of variable extent of endosome acidification occurs in ligand-receptor endosomes, where pH values range from 6.4 to 5.0 depending on the receptor involved¹²⁻¹⁵. Although *Toxoplasma* show a specific attachment phase before entry²⁻⁴, neither the receptor involved in this step nor the distribution of other cell-surface receptors within the phagosome have been characterized.

The significance of phagosome acidification to intracellular *Toxoplasma* is paramount to survival, judging from their high susceptibility to mild acidic conditions when they are extracellular. Thus, the failure of acidification exhibited by live *Toxoplasma*-containing vacuoles in normal macrophages enables *Toxoplasma* to survive within cells normally considered microbicidal. In support of this hypothesis, two other intracellular parasites which resist lysosome fusion have been shown to reside in neutral (pH = 6.9)⁷ or weakly acidic (pH 6.1)²⁴ compartments in normal macrophages. Increased capacity for phagosome acidification in activated macrophages and in normal macrophages in the presence of specific antibody probably contributes to the destruction of intracellular *Toxoplasma*.

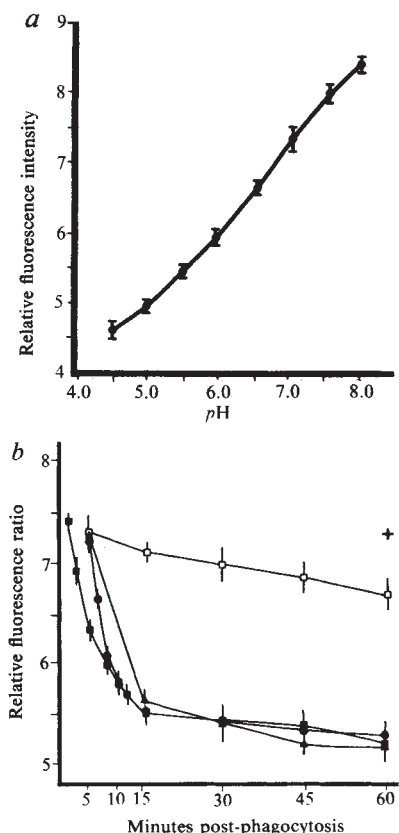


Fig. 3 *a*, In situ standard curve relating fluorescence intensity (I_0) ratio of intracellular FITC-labelled *Toxoplasma* to fixed pH values. *b*, Fluorescence intensity ratios from macrophage monolayers containing fluorescein-labelled *Toxoplasma* vacuoles at intervals after phagocytosis. A decrease in I_0 represents a shift in the intravacuolar pH to acidic conditions (*a*). A slight decline in I_0 is evident during the first hour after infection from modified live *Toxoplasma*-containing vacuoles in normal macrophages ($\square$), probably due to the 20% of *Toxoplasma* in this preparation which fuse with acidic vesicles. In contrast, a rapid decrease in I_0 is observed from macrophage vacuoles containing heat-killed *Toxoplasma* and fluorescein dextran ($\blacktriangle$), and FITC-antibody-labelled live *Toxoplasma* ($\blacksquare$) in normal macrophages and from live *Toxoplasma* and fluorescein dextran ($\bullet$) in activated macrophages.

Methods. *a*, Macrophage monolayers containing FITC-*Toxoplasma* (see Fig. 2*a, b*) were incubated in sodium citrate (pH 4.0, 4.5, 5.0) or PBS (pH 5.5, 6.0, 6.5, 7.0, 7.5, 8.0) containing 10 μ M monensin to collapse cell pH gradients and clamp intracellular compartments at the fixed pH of the external buffer. Fluorescence intensity was recorded from macrophages attached to coverslips placed in 1-cm cuvettes using an Aminco-Bowman spectrofluorometer with 10-nm bandwidth excitation (altered between 450 and 495) and 5-nm bandwidth emission centred at 520 nm, $n=6$ (ref. 21). Correction for background scatter and autofluorescence was made by subtraction of I_0 readings from a similar set of monolayers which did not contain fluorescein. *b*, Infected macrophages labelled with fluorescein (cell density 2×10^6 per coverslip; 2–5 *Toxoplasma* per host cell) were rinsed in PBS and I_0 recorded in PBS at 37°C. Corrected values were obtained by subtracting background fluorescence readings from unlabelled infected monolayers. A separate set of three coverslips was used to determine mean I_0 ratio for each time point. Intracellular acidification of heat-killed or antibody-coated *Toxoplasma*-containing vacuoles was reversible by addition of 10 μ M monensin in PBS at pH 7.2 (+), but was not affected by changes in external buffer pH in the absence of ionophore.

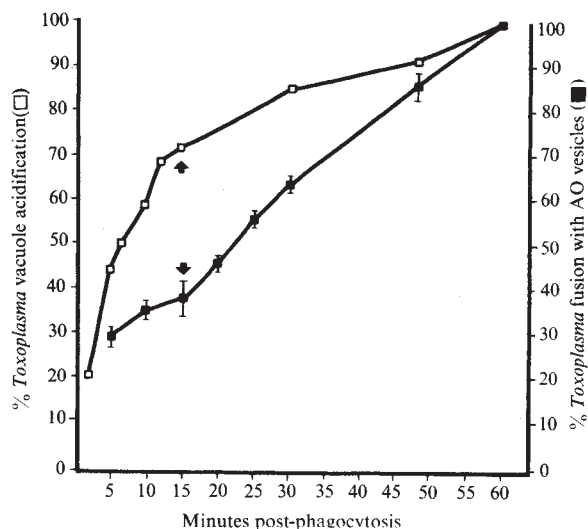


Fig. 4 Comparison of the rate of fusion of acidic vesicle (labelled with acridine orange, AO) ($\blacksquare$) with pH changes ($\square$) in normal macrophage containing heat-killed or antibody-coated *Toxoplasma*. Most acidification occurs before the fusion of acidic vesicles including lysosomes. The % of total acidification is calculated from I_0 data points in Fig. 3*b* representing FITC-labelled *Toxoplasma* ingested during a 1-min pulse. The extent of acidic vesicle fusion was obtained from normal macrophages labelled with acridine orange and challenged similarly with FITC-*Toxoplasma*. Three separate coverslips were used to score each time point. Similar kinetic curves for fusion rate were obtained with heat-killed *Toxoplasma* in normal macrophages and live *Toxoplasma* in activated macrophages.

We thank Dr Jack Remington for providing *Toxoplasma* strains and Dr Elmer Pfefferkorn for fibroblast cultures. This work was supported in part by a Grant-in-Aid of Research from Sigma Xi, the Scientific Research Society and NIAID research grant AI22442-01.

Received 9 October 1984; accepted 20 March 1985.

- Goren, M. B. A. *Rev. Microbiol.* **31**, 507-533 (1977).
- Jones, T. C., Yeh, S. & Hirsch, G. J. *exp. Med.* **136**, 1157-1172 (1972).
- Aikawa, M., Komata, Y., Asai, T. & Midorikawa, O. *Am. J. Path.* **87**, 285-296 (1977).
- Jones, T. C. & Hirsch, J. G. J. *exp. Med.* **136**, 1173-1194 (1972).
- Allen, R. C. in *Lysosomes in Biology and Pathology* Vol. 6 (eds Dingle, J. T., Jacques, P. J. & Shaw, I. H.) 197-233 (North-Holland, Amsterdam, 1979).
- Geisow, M. J., Hart, P. D. & Young, M. R. *J. Cell Biol.* **89**, 645-652 (1981).
- Weidner, E. & Sibley, L. D. *J. Protozool.* **32**, 311-317 (1985).
- Meneil, P. L., Tanasugarn, L., Meigs, J. B. & Taylor, D. L. *J. Cell Biol.* **97**, 692-702 (1983).
- Galloway, C. J., Dean, G. E., Marsh, M., Rudnick, G. & Mellman, I. *Proc. natn. Acad. Sci. U.S.A.* **80**, 3334-3338 (1983).
- Forgac, M. & Cantley, L. *J. biol. Chem.* **259**, 8101-8105 (1984).
- Xie, X. S., Stone, D. K. & Racker, E. *J. biol. Chem.* **258**, 14834-14838 (1982).
- Yamashiro, D. J., Fluss, S. R. & Maxfield, F. R. *J. Cell Biol.* **97**, 929-934 (1983).
- Tycko, B. & Maxfield, F. R. *Cell* **28**, 643-651 (1982).
- Tietze, C., Schlesinger, P. & Stahl, P. *J. Cell Biol.* **92**, 417-424 (1982).
- Yamashiro, D. J., Tycko, B., Fluss, S. R. & Maxfield, F. R. *Cell* **37**, 789-800 (1984).
- Murray, H. W. & Cohn, Z. A. *J. exp. Med.* **150**, 938-949 (1979).
- Nathan, C. F. *Trans. R. Soc. trop. Med. Hyg.* **77**, 620-630 (1983).
- Wilson, C. B., Tsai, V. & Remington, J. S. *J. exp. Med.* **151**, 328-346 (1980).
- Allison, A. C. & Young, M. R. in *Lysosomes in Biology and Pathology* Vol. 1 (eds Dingle, J. T. & Feil, H. B.) 600-628 (North-Holland, Amsterdam, 1969).
- Pfefferkorn, E. R. & Pfefferkorn, L. C. *Expl Parasit.* **39**, 365-376 (1976).
- Ohkuma, S. & Poole, B. *Proc. natn. Acad. Sci. U.S.A.* **75**, 3327-3331 (1978).
- Heiple, J. M. & Taylor, D. L. *J. Cell Biol.* **94**, 143-149 (1982).
- Allen, R. D. & Fok, A. K. *J. Cell Biol.* **97**, 566-570 (1983).
- Horwitz, M. A. & Maxfield, F. R. *J. Cell Biol.* **99**, 1936-1943 (1984).
- Khavkin, T. N., Freidlin, I. S. & Shustrov, A. K. *Acta microbiol. hung.* **27**, 9-21 (1980).

Mitogens trigger a calcium-independent signal for proliferation in phorbol-ester-treated lymphocytes

Erwin W. Gelfand, Roy K. Cheung, Gordon B. Mills & Sergio Grinstein

Divisions of Immunology/Rheumatology and Cell Biology, Research Institute, Hospital for Sick Children, Toronto, Canada M5G 1X8

The activation of T lymphocytes by mitogens requires at least two signals; the first, delivered to T cells by a mitogen in conjunction with accessory cells (monocytes/macrophages), leads to the generation of the second signal, interleukin-2 (IL-2). The first signal also induces the expression of IL-2 receptors on the surface of a subpopulation of T cells; binding of IL-2 to its receptor then initiates a cascade of events culminating in DNA synthesis by these cells¹⁻⁴. Certain compounds act synergistically with mitogens in promoting T-cell proliferation by substituting for the activities of interacting cells or their products. For example, the phorbol ester 12-O-tetradecanoyl phorbol 13-acetate (TPA) has been shown to restore the ability of macrophage-depleted T-cell populations to respond to mitogenic lectins⁵⁻⁷. Transmembrane fluxes of calcium, leading to increased free cytosolic calcium concentrations ($[Ca^{2+}]_i$), have been demonstrated following mitogen binding to lymphocytes and have been implicated in the initiation of cell proliferation⁸⁻¹⁰. We show here that the effect of TPA on lymphocyte proliferation occurs in the absence of extracellular Ca^{2+} or detectable changes in $[Ca^{2+}]_i$, but only in the presence of mitogens. This suggests that in cells which have been incubated with the phorbol ester, mitogens can induce proliferation by a calcium-independent signal.

Mitogenic lectins such as phytohaemagglutinin (PHA) and concanavalin A, and non-mitogenic lectins such as wheat germ agglutinin, lead to a rapid rise in $[Ca^{2+}]_i$ even when there is no proliferative response⁹, and this hampers attempts to evaluate the requirement for an increase in $[Ca^{2+}]_i$ in the initiation of cell proliferation. However, we have shown previously that, in contrast to PHA-induced activation of monocyte-depleted human T cells, the T-cell mitogen¹¹ protein A (SpA from

Table 1 Mitogen-dependent effects on levels of cytoplasmic Ca^{2+} and lymphocyte proliferation

	$[Ca^{2+}]_i$ (nM)				³ H-thymidine incorporation (c.p.m.)
	0	10 min	3 h	24 h	
PBM + PHA	108 ± 2.0	235 ± 8.0	187 ± 7.4	114 ± 3.1	44,090 ± 2,670
PBM + SpA	110 ± 2.7	128 ± 12.4	180 ± 5.5	114 ± 3.8	22,165 ± 1,860
E ⁺ + PHA	106 ± 2.0	245 ± 7.8	186 ± 2.4	116 ± 3.5	62,210 ± 2,110
E ⁺ + SpA	106 ± 2.4	110 ± 5.3	113 ± 3.4	114 ± 3.0	4,520 ± 760

Peripheral blood mononuclear cells (PBM) were isolated following Ficoll-Hypaque gradient centrifugation. Erythrocyte-rosetting T cells (E⁺) were obtained by rosette enrichment as described elsewhere²³. Cells ($200 \mu\text{l}$, $5 \times 10^5 \text{ ml}^{-1}$) were cultured with PHA ($10 \mu\text{g ml}^{-1}$, Difco) or SpA ($100 \mu\text{g ml}^{-1}$, Pharmacia) for 72 h in RPMI-1640 containing 10% fetal calf serum. Four hours before the end of the culture, $1 \mu\text{Ci } ^3\text{H-thymidine}$ (6.7 Ci mmol^{-1}) was added; labelled DNA was counted in a liquid scintillation counter. The mitogens were added at time 0 and $[Ca^{2+}]_i$ levels recorded at 10 min, 3 h and 24 h. For monitoring $[Ca^{2+}]_i$, $10 \times 10^6 \text{ cells ml}^{-1}$ were incubated with $10 \mu\text{M}$ acetoxymethyl quin-2 (generously provided by Dr T. J. Rink, and prepared as a 2 mM stock solution in dimethyl sulphoxide) as described previously¹⁰. The cells were incubated with quin-2 before each assay point. A Perkin-Elmer spectrofluorimeter (Model 650-40) was used for fluorescence measurements, for which the cells were suspended in HEPES-buffered saline containing 1 mM Ca^{2+} (ref. 10). Calculation of free cytosolic Ca^{2+} concentration ($[Ca^{2+}]_i$) was as described previously^{10,24}. Although reported to affect the response of lymphocytes to mitogen⁹, the concentration of quin-2 used here was shown to be non-mitogenic and did not significantly interfere with mitogen-induced proliferation. Results represent the means $\pm$ s.e.m. of three separate experiments carried out in duplicate (Ca^{2+} determinations) or triplicate ($^3\text{H-thymidine}$ incorporation). In the absence of mitogen, background c.p.m. were $<1,000$ and $[Ca^{2+}]_i$ was $106-112 \text{ nM}$.

Table 2 TPA modulation of mitogen-induced changes in levels of cytosolic free Ca^{2+} and lymphocyte proliferation

	$[Ca^{2+}]_i$ (nM)				³ H-thymidine incorporation (c.p.m.)
	0	10 min	3 h	24 h	
TPA (16 h)	112 ± 3.0	112 ± 3.2	110 ± 2.5	112 ± 2.5	3,100 ± 270
PHA	111 ± 2.3	236 ± 5.6	191 ± 5.3	—	40,530 ± 1,350
TPA (16 h) → PHA	108 ± 1.2	246 ± 7.6	204 ± 4.2	—	52,400 ± 5,790
PHA	114 ± 3.2	260 ± 6.2	201 ± 3.2	118 ± 4.6	34,870 ± 2,480
TPA (1 h) → PHA	108 ± 3.2	256 ± 8.2	202 ± 3.5	120 ± 5.2	47,980 ± 2,870
SpA	109 ± 2.4	112 ± 1.6	112 ± 4.0	—	4,040 ± 670
TPA (16 h) → SpA	112 ± 1.5	120 ± 5.6	115 ± 5.1	—	23,110 ± 3,160
SpA	105 ± 3.2	110 ± 6.4	116 ± 6.2	114 ± 5.2	2,250 ± 690
TPA (1 h) → SpA	116 ± 2.1	118 ± 2.6	120 ± 3.6	122 ± 5.2	20,570 ± 1,160

Human E-rosetting T cells ($10 \times 10^6 \text{ ml}^{-1}$) were incubated with TPA (10^{-9} M) (Sigma) for 1 or 16 h in RPMI-1640 plus 10% fetal calf serum. The cells were then washed three times in RPMI containing defatted bovine serum albumin (BSA, Sigma) and twice with RPMI to remove residual TPA²⁵. Aliquots of the cells were then either loaded with quin-2 for $[Ca^{2+}]_i$ determinations or used in studies of lymphocyte proliferation (see Table 1). Results represent the means $\pm$ s.e.m. of three separate experiments.

Staphylococcus aureus, Cowan I strain) does not produce changes in $[Ca^{2+}]_i$ or DNA synthesis in the absence of monocytes¹². Monocytes preincubated with SpA trigger a rapid increase in $[Ca^{2+}]_i$ as well as a proliferative response in these isolated T cells. We therefore used this system to determine whether TPA can serve as a co-mitogen and whether the effect of TPA on SpA-induced proliferation of isolated human T cells (in the absence of monocytes) requires, or is accompanied by, changes in $[Ca^{2+}]_i$.

Table 1 illustrates the effects of PHA and SpA on $[Ca^{2+}]_i$ monitored by quin 2 fluorescence during the first 24 h and by the proliferative response measured using $^3\text{H-thymidine}$ incorporation at 72 h. Incubation of a mixed population of peripheral blood mononuclear cells (PBM) with PHA or SpA resulted in significant incorporation of $^3\text{H-thymidine}$ and increases in $[Ca^{2+}]_i$. The response to PHA was rapid, $[Ca^{2+}]_i$ more than doubling within 10 min, then falling progressively and reaching

Table 3 Mitogen-dependent changes observed in the presence of EGTA

Expt a		³ H-thymidine incorporation (c.p.m.)
Initial incubation	Second incubation	
Medium	Medium	1,100 ± 200
Medium	TPA	2,100 ± 230
Medium	PHA	22,840 ± 3,240
Medium	TPA → PHA	25,110 ± 2,910
EGTA	TPA	2,190 ± 270
EGTA	PHA	4,220 ± 730
EGTA	TPA → PHA	24,790 ± 2,720
Medium	SpA	4,570 ± 660
Medium	TPA → SpA	16,200 ± 570
EGTA	SpA	5,870 ± 320
EGTA	TPA → SpA	18,500 ± 1,080

Expt b		[Ca ²⁺] _i (nM)	
Initial incubation	Second incubation		
Medium	106 ± 2.0	PHA	245 ± 7.8
Medium	112 ± 3.0	TPA	112 ± 3.2
TPA	108 ± 3.2	PHA	256 ± 8.2
EGTA	102 ± 6.6	TPA	103 ± 5.2
EGTA	105 ± 3.0	PHA	107 ± 2.8
EGTA → TPA	108 ± 2.6	PHA	108 ± 3.2

E-rosetting T cells were used in these studies. Expt a: Determination of ³H-thymidine incorporation. In the initial incubation, 5×10^5 cells ml⁻¹ were incubated with medium containing EGTA (2 mM) or medium (RPMI-1640) alone for 30 min at 37 °C. In the second incubation, TPA (10^{-9} M) was added where indicated for 1 h, followed by the addition of mitogen for 1 h at 37 °C. Only then were the cells washed three times with RPMI-containing defatted BSA and twice in RPMI. Aliquots (200 µl) of cells (5×10^5 ml⁻¹) were then incubated for 72 h in RPMI-1640 plus 10% fetal calf serum. Previous studies have shown that a 1-h incubation with PHA, followed by cell washing, results in 80–90% of the proliferative response observed when PHA is left in culture for the entire 72 h¹⁶. Results represent the means ± s.e.m. of three experiments carried out in triplicate. Expt b: Determination of [Ca²⁺]_i levels. Initial incubation: cells were first incubated with quin-2 (10 µM) for 60 min, washed, then incubated with medium containing EGTA (2 mM) or medium only (see Table 1) for 30 min at 37 °C. TPA (10^{-9} M) was then added for 1 h as indicated. Baseline levels of [Ca²⁺]_i were obtained. In the second incubation, PHA (or TPA) was added and [Ca²⁺]_i levels determined 10 min later. Results represent the means ± s.e.m. of three separate experiments carried out in duplicate.

baseline levels after 24 h. The response to SpA was somewhat slower, with a period of 2–3 h required to reach maximum [Ca²⁺]_i levels. This delay probably reflects the requirement for monocyte/T-cell interaction in the response to SpA¹². Purified T cells (E⁺) respond to PHA in a similar manner to the unfractionated PBM, with both a rapid increase in [Ca²⁺]_i and a greater proliferative response. In contrast, as previously reported¹², incubation of purified T cells with SpA resulted in a greatly reduced ³H-thymidine incorporation and no significant change in [Ca²⁺]_i.

TPA is the most potent of the group of phorbol diesters known to promote tumours or induce cellular differentiation¹³. By itself, TPA (10^{-9} M) is weakly mitogenic but does not lead to changes in [Ca²⁺]_i (ref. 8 and Table 2). We found that preincubation of cells with TPA for 1 or 16 h, followed by the addition of PHA, produced a greater proliferative response than observed with cells not preincubated with TPA. The increase in cytosolic free Ca²⁺ following addition of PHA was similar with or without TPA pretreatment.

When purified T cells were incubated with TPA for 1 or 16 h, they subsequently exhibited a proliferative response to SpA (Table 2); the magnitude of ³H-thymidine incorporation was similar to that seen with PBM (Table 1). In the presence of SpA, [Ca²⁺]_i levels were only marginally higher in cells preincubated with TPA than in cells exposed to SpA alone; monitoring of [Ca²⁺]_i indicated little change in levels during the 24-h period following addition of SpA. These results indicate that preincubation enables T cells to proliferate in response to SpA without an accompanying rise in [Ca²⁺]_i.

In monocyte-dependent systems, TPA can often restore the proliferative response to mitogens in monocyte-depleted cell

populations^{4–6}. Although TPA can augment the production of interleukin-1 (a monocyte-derived lymphokine)^{14,15} the effects of TPA described in Table 2 are monocyte-independent and the drug seems to act directly on purified T cells. These effects of TPA on SpA-induced T-cell proliferation can be easily distinguished from those observed following addition of monocytes; in the presence of monocytes, [Ca²⁺]_i increased significantly (Table 1, PBM) whereas with TPA pretreatment of purified T cells there was little deviation from baseline [Ca²⁺]_i following the addition of SpA (Table 2).

We have shown previously that triggering of a PHA-induced proliferative response is dependent on the presence of extracellular Ca²⁺ (ref. 16). We therefore investigated whether TPA treatment would produce a PHA-induced proliferative response in the absence of raised [Ca²⁺]_i. We found that, in contrast to cells in Ca²⁺-containing medium, T lymphocytes treated with the Ca²⁺ chelator EGTA and then incubated with PHA for 1 h in EGTA-containing medium, failed to incorporate ³H-thymidine when returned to normal (and lectin-free) medium for the remaining 72 h (Table 3, expt a and ref. 16). Cells treated similarly, except for preincubation with TPA, demonstrated a response to PHA (and SpA). As the TPA preincubation was carried out with EGTA-treated cells in EGTA-containing medium, this result suggests that the TPA effects do not require extracellular Ca²⁺, in contrast to previous reports¹⁷. Cells pre-treated with EGTA and incubated with TPA and PHA in EGTA-containing medium showed no increase in [Ca²⁺]_i (Table 3, expt b), but did show an increase in ³H-thymidine incorporation (Table 3, expt a).

The importance of increases in Ca²⁺ as a signal for mitogen-induced proliferation has been demonstrated in several different systems (reviewed in ref. 18), although the nature of the Ca²⁺ signal and the target for its activity are unknown. On the other hand, TPA is known to be co-mitogenic for lymphocytes and may act by directly activating protein kinase C, leading to protein phosphorylation^{19–22}. The present report, as well as previous studies⁸, have shown that TPA does not mobilize Ca²⁺ as measured directly using the fluorescent indicator quin-2. We used two independent systems, one in which changes in [Ca²⁺]_i do not occur in response to mitogen alone and one in which changes in [Ca²⁺]_i were prevented by the presence of EGTA. In both systems, following treatment with TPA, T lymphocytes become able to proliferate in response to mitogens. As TPA, an activator of protein kinase C, is itself non-mitogenic, this suggests that another, calcium-independent, signal can be generated by the mitogen; the nature of this signal remains to be elucidated.

This work was supported by the MRC of Canada and the National Cancer Institute.

Received 8 August 1984; accepted 19 April 1985.

- Palacios, R. *Immun. Rev.* **63**, 73–110 (1982).
- Mizel, S. B. *Immun. Rev.* **63**, 51–72 (1982).
- Smith, K. A., Wang, H.-M. & Cantrell, D. A. in *Progress in Immunology* (eds Yamamura, Y. & Tada, T.) 259–271 (Academic, New York, 1983).
- Gillis, S. *J. clin. Immunol.* **3**, 1–13 (1983).
- Rosenreich, D. L. & Mizel, S. B. *J. Immunol.* **123**, 1749–1754 (1979).
- Farrar, J. J., Mizel, S. B., Fuller-Bonar, J., Farrar, W. L. & Hilfiker, M. L. *J. Immunol.* **125**, 793–798 (1980).
- Koretzky, G. A., Daniele, R. P. & Nowell, P. C. *J. Immunol.* **128**, 1776–1780 (1982).
- Tsien, R. Y., Pozzan, T. & Rink, T. J. *Nature* **295**, 68–71 (1982).
- Hesketh, T. R., Smith, G. A., Moore, J. P., Taylor, M. V. & Metcalfe, J. C. *J. biol. Chem.* **258**, 4876–4882 (1983).
- Gelfand, E. W., Cheung, R. K. & Grinstein, S. *J. cell. Physiol.* **121**, 533–539.
- Schuurman, R. K. B., Gelfand, E. W. & Dosch, H. M. *J. Immunol.* **125**, 820–826 (1980).
- Lederman, H. M., Lee, J. W. W., Cheung, R. K., Grinstein, S. & Gelfand, E. W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6827–6830.
- Diamond, L., O'Brien, T. G. & Rovera, G. *Life Sci.* **23**, 1979–1988 (1978).
- Mizel, S. B., Rosenreich, D. & Oppenheim, J. J. *Cell Immunol.* **40**, 230–235 (1978).
- Krakauer, T., Mizel, D. & Oppenheim, J. J. *J. Immunol.* **129**, 939–941 (1982).
- Cheung, R. K., Grinstein, S. & Gelfand, E. W. *J. Immunol.* **131**, 2291–2295 (1983).
- Mastro, A. M. & Smith, M. C. *J. cell. Physiol.* **116**, 51–56 (1983).
- Lichtman, A. H., Segel, G. B. & Lichtman, M. A. *Blood* **61**, 413–422 (1983).
- Nishizuka, Y. *Nature* **308**, 693–698 (1984).
- Berridge, M. J. *Biochem. J.* **220**, 345–360 (1984).
- Kikkawa, W., Takai, Y., Tanaka, Y., Miyake, R. & Nishizuka, Y. *J. biol. Chem.* **258**, 11442–11445 (1983).
- Vandenbark, G. R., Kuhn, L. J. & Nield, J. E. *J. clin. Invest.* **73**, 448–457 (1984).
- Cheung, R. K., Grinstein, S. & Gelfand, E. W. *J. clin. Invest.* **70**, 632–638 (1982).
- Tsien, R. Y., Pozzan, T. & Rink, T. J. *J. Cell Biol.* **94**, 325–334 (1982).
- Dicker, P. & Rozengurt, E. *J. cell. Physiol.* **109**, 99–109 (1981).

Sequence of reovirus haemagglutinin predicts a coiled-coil structure

Rhonda Bassel-Duby^{*†}, Anula Jayasuriya[†],
Devjani Chatterjee[‡], Nahum Sonenberg^{*},
Jacob V. Maizel Jr[‡] & Bernard N. Fields^{†§}

^{*} Department of Biochemistry, McGill University, Montréal, Québec, Canada H3G 1Y6

[†] Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

[‡] Section of Molecular Structure, Laboratory of Molecular Genetics, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20205, USA

[§] Shipley Institute of Medicine and Department of Medicine, Division of Infectious Diseases, Brigham and Women's Hospital, Boston, Massachusetts 02115, USA

The use of modern techniques has led to new insights into the molecular mechanisms of viral pathogenesis¹. Although the infectious process is quite complex, it is clear that one critical stage, the interaction of viral attachment proteins with cell-surface receptors, often has a major role in determining the pattern of infection². The mammalian reoviruses have served as useful models for understanding the molecular basis of viral pathogenesis³. The mammalian reovirus haemagglutinin ($\sigma 1$ protein), which is an outer capsid protein, has been shown to be a major factor in determining virus-host cell interactions⁴. To further our understanding of the structure and function of the haemagglutinin, we have cloned a complementary DNA copy of the reovirus type 3 *S1* double-stranded RNA gene which encodes the virus haemagglutinin and have sequenced the DNA complementary to the *S1* gene. Analysis of the predicted amino-acid sequence of the virus haemagglutinin has allowed us to determine that the amino-terminal portion contains an α -helical coiled-coil structure and that the carboxy-terminal portion contains the receptor-interacting domains. Using this information, we propose here a model of how the reovirus haemagglutinin is attached to the virus particle.

Fig. 1 The determined nucleotide sequence of reovirus type 3 *S1* gene and predicted amino-acid sequences for both open reading frames. The one-letter amino-acid notation used is recommended by the IUPAC-IUB Commission on Biochemical Nomenclature²². The numbers given above the amino-acid sequence designate the residue positions for the first open reading frame. The numbers on the right show nucleotide positions. The α -helical region is indicated by a thick line above the amino-acid residues. The potential glycosylation sites are marked by an asterisk. The mutation site for an immune-selected variant (K) is indicated by an arrow.

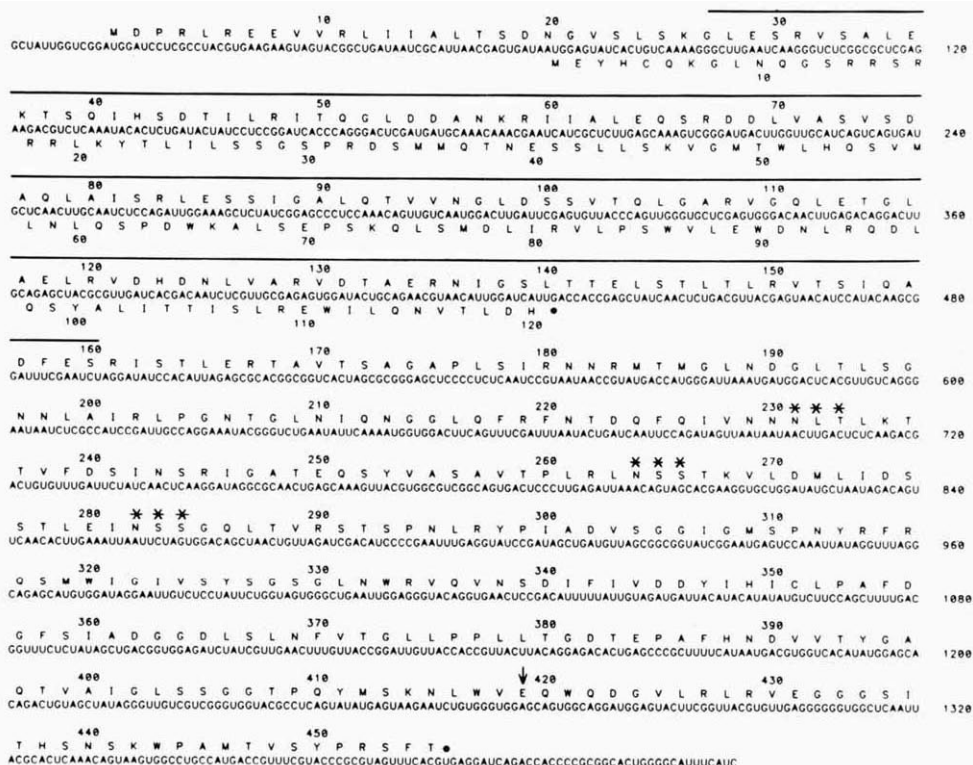
Methods. Genomic double-stranded (ds) RNA was isolated from purified reovirus type 3 (Dearing) and, using the enzyme poly(A) polymerase (BRL), ~30 AMP molecules were added to the 3' termini²³. The polyadenylated RNA was used as a template to synthesize oligo(dT)-primed single-stranded (ss) cDNA²⁴. The complementary ss cDNA fragments were hybridized to obtain ds cDNA fragments which were inserted into pBR322 by G-C tailing²⁵. The recombinant plasmids were used to transform *Escherichia coli* MC1061 cells²⁴. To obtain an *S1* gene probe, reovirus dsRNA was tailed with radiolabelled cytidine 3',5'-bisphosphate using *T₄* RNA ligase²⁶. The radiolabelled *S1* dsRNA was then used to screen the cDNA library. A full-length cDNA clone of the *S1* gene was isolated. Restriction enzyme fragments for this cDNA clone were ligated into bacteriophage vector M13 mp9 RF-DNA²⁷ and sequenced by the dideoxy-chain terminator method²⁸.

Figure 1 shows the nucleotide sequence of the entire reovirus type 3 *S1* gene. The gene consists of 1,416 nucleotide bases and has two open reading frames. The first open reading frame begins at nucleotide 13 and terminates at nucleotide 1,377, encoding a protein of 455 amino acids and computed relative molecular mass of 49,071. This protein corresponds to the reovirus haemagglutinin ($\sigma 1$ protein) as determined by the previously estimated sizes^{3,4}.

The predicted amino-acid sequence of the virus haemagglutinin contains three potential glycosylation sites (see Fig. 1). Although it has been reported that the haemagglutinin of reovirus grown in the L strain of mouse fibroblasts can incorporate radiolabelled amino-sugars at a low level⁵, there has been no confirmation that the three potential sites are glycosylated in the haemagglutinin.

A second open reading frame in the *S1* gene begins at nucleotide 71 and terminates at nucleotide 430 and would encode a polypeptide of 120 amino acids. The presence of two open reading frames agrees with both ribosome binding studies⁶ and *in vitro* translation studies⁷ which conclude that the *S1* gene has two initiation sites. Recently, the 120-amino-acid protein has been detected in reovirus-infected cells⁸.

Examinations of the amino-acid sequence of the reovirus type 3 haemagglutinin reveals certain features, in particular, a heptapeptide repeat pattern (*a-b-c-d-e-f-g*) between amino acids 28 and 158, representing approximately one-third of the molecule at the amino-terminal end. In this pattern, amino acids at positions *a* and *d* are primarily hydrophobic (Table 1). Another striking feature of the sequence is that there are two proline residues flanking this heptapeptide repeat at amino acid positions 3 and 176, but no prolines within the heptapeptide repeat pattern. The heptapeptide repeat, the position of the hydrophobic amino acid and the absence of both proline and aromatic amino acids from the heptapeptide repeat are typical of an α -helical coiled-coil structure where the hydrophobic residues at positions *a* and *d* form the interface between the α -helices^{9,10}; this agrees with computer analysis of protein secondary-structure predictions¹¹ which shows this area to



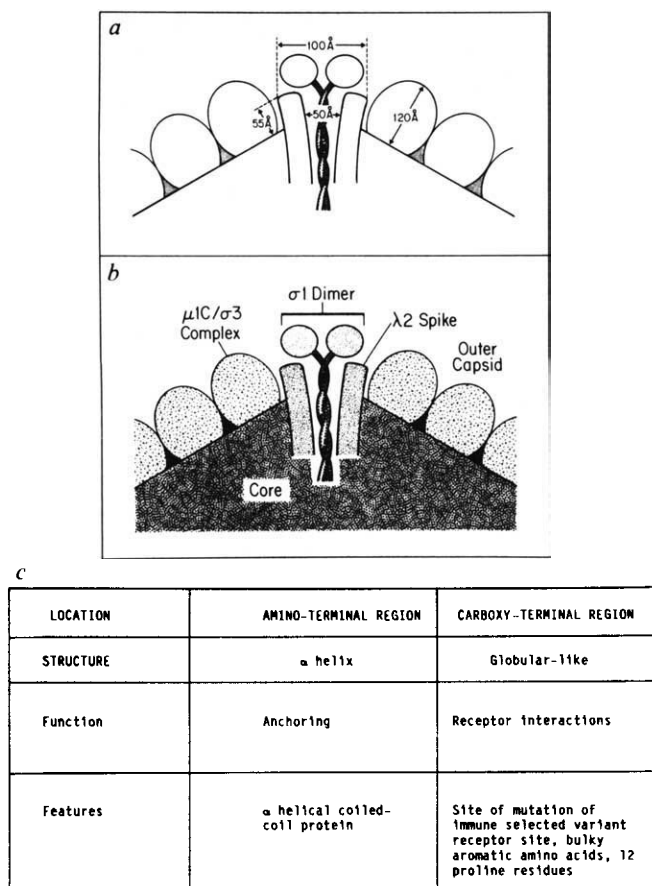


Fig. 2 Schematic representation of the morphology of the outer capsid of reovirus type 3. *a*, Dimensions of the virion. *b*, Orientation of the $\sigma 1$ protein in the virus. The α -helical region is shown to extend through the $\lambda 2$ channel into the viral core. The globular-like structure sits on top of the $\lambda 2$ channel and interacts with host cell receptors. *c*, Summary of structure-function of the virus haemagglutinin.

consist of an α -helical structure. Similar predictions of a coiled-coil structure based on sequence data and computer analysis were reported for the influenza virus haemagglutinin¹². This structure was subsequently verified by crystallography and shown to represent ~23% of the polypeptide that provides the major force stabilizing the trimeric subunit interactions¹³.

By definition, the α -helices of a coiled-coil structure are wound around each other; this predicts at least a dimer form for the reovirus haemagglutinin. Studies show that there are 24 haemagglutinin molecules per virion which are located on the 12 projections/spikes on the outer capsid¹⁴. Thus, both the stoichiometry and the coiled-coil structure support a dimer form for the haemagglutinin.

In contrast to the amino-terminal portion, the carboxy-terminal portion of the virus haemagglutinin shows no dominant structural pattern. Computer analysis of protein secondary-structure predictions¹¹ show that this area is a mixture of short α -helices, β -sheets and random coils and that it contains many turns. This area also contains 12 proline residues and bulky aromatic amino acids, unlike the α -helical amino-terminal portion. All these features suggest a globular-like structure at the carboxy-terminal portion of the haemagglutinin protein.

Recent studies on a reovirus variant strain (designated variant K)¹⁵ have localized the receptor-binding region to the carboxy terminus of the virus haemagglutinin (R.B.-D., unpublished data). In particular, the *S1* gene of this immune-selected variant, in which the receptor-binding site is altered¹⁶, is similar to the *S1* gene of reovirus type 3; the only nucleotide base change in the variant *S1* gene is a point mutation at nucleotide 1,267

Table 1 Heptapeptide repeat pattern for the amino-acid sequence of reovirus type 3 haemagglutinin

a	b	c	d	e	f	g	Amino-acid no.
Val	Ser	Ala	Leu	Glu	Ser	Arg	28-31
Ser	Gln	Ile	Leu	Glu	Lys	Thr	32-38
Ile	Leu	Arg	Ile	Thr	Ser	Thr	39-45
Leu	Asp	Asp	Ala	Asn	Lys	Arg	46-52
Ile	Ile	Ala	Leu	Glu	Gln	Ser	53-59
Arg	Asp	Asp	Leu	Val	Ala	Ser	60-66
Val	Ser	Asp	Ala	Gln	Leu	Ala	67-73
Ile	Ser	Arg	Leu	Glu	Ser	Ser	74-80
Ile	Gly	Arg	Leu	Gln	Thr	Val	81-87
Val	Asn	Gly	Leu	Asp	Ser	Ser	88-94
Val	Thr	Gln	Leu	Gly	Ala	Arg	95-101
Val	Gly	Gln	Leu	Glu	Thr	Gly	102-108
Leu	Ala	Glu	Leu	Arg	Val	Asp	109-115
His	Asp	Asn	Leu	Leu	Ala	Arg	116-122
Val	Asp	Thr	Ala	Glu	Arg	Asn	123-129
Ile	Gly	Ser	Leu	Thr	Thr	Glu	130-136
Leu	Ser	Thr	Leu	Thr	Leu	Arg	137-143
Val	Thr	Ser	Ile	Gln	Ala	Asp	144-150
Phe							151-157
16/19	3/18	4/18	18/19	2/19	6/19	2/19	158
89	17	22	99	10	34	10	No. of hydrophobic amino acids
							% Hydrophobic amino acids

(corresponding to a change in amino acid 419 from glutamic acid to lysine). Thus, the receptor-binding site is in the carboxy-terminal portion of the haemagglutinin protein.

The protein sequence of the reovirus haemagglutinin was compared with all the protein sequences available in the National Biomedical Research Foundation database, using the Goad-Kanehisa program for finding local homologies¹⁷. Significantly high scores (8 and 7 s.d.) were obtained with nematode myosin and rabbit tropomyosin. Optimal alignment of reovirus haemagglutinin (amino acids 27-100) and nematode myosin (amino acids 1,391-1,464) gives an average of 28% identical residues. When comparisons are based on grouping amino acids containing similar side chains, the extent of homology increases to an average of 51%. However, myosin¹⁸ and tropomyosin¹⁰ both consist of α -helical coiled-coil structures. To evaluate whether the observed homologies were due to evolutionary relatedness or structural relatedness, further computer analysis was performed on the homologies and revealed scores (4.5 and 4 s.d.) which indicated definite structural relatedness between reovirus haemagglutinin and myosin/tropomyosin. Because of the lack of any known functional similarities, the evolutionary relatedness of these proteins remains uncertain.

Figure 2 shows a model of the virus haemagglutinin and its orientation in the virus. Also shown are the known dimensions of the outer capsid, the location of the $\mu 1C/\sigma 3$ complex, the $\lambda 2$ projections/spikes and the channel within the $\lambda 2$ protein. Evidence that the haemagglutinin lies close to the $\lambda 2$ protein comes from analysis of extragenic suppressor mutations¹⁹. Temperature-sensitive (ts) mutants of reovirus containing defects in the $\lambda 2$ protein (tsB-352) revert to a non-ts phenotype with second mutations in the $\sigma 1$ protein. Such second-site mutations occur predominantly in proteins which share an intimate structural relationship. Further evidence for a close interaction between $\lambda 2$ and $\sigma 1$ proteins is the fact that although $\sigma 1$ protein is the major neutralization antigen, monoclonal antibodies against $\lambda 2$ protein also result in neutralization²⁰, presumably through steric hindrance of the $\sigma 1$ protein function.

The $\sigma 1$ protein has been drawn in the model as a dimer containing an extended α -helix whose length when fully extended is ≈ 180 Å. Unlike the adenovirus fibre protein, which projects as an extended rod from the surface of the adenovirus core, no structure has been seen to project from the surface of

the reovirus particle. Therefore, the extended α -helix of the protein must be accommodated within the virion. The location of the haemagglutinin over the $\lambda 2$ channel makes it likely that the extended α -helical portion of the haemagglutinin extends into the channel, while the globular portion containing the receptor-interacting region is present on the outer surface of the virus particle. This model is consistent with the known channel diameter of the core spike (50 Å) and the radius of the outer capsid (120 Å)²¹. The location of the α -helix in the $\lambda 2$ channel would provide an anchor, attaching the $\sigma 1$ protein to the virus particle. Recent work in this laboratory on the morphology of the reovirion supports the proposed model (M. Nibert, unpublished data). The present studies should now make it possible to analyse, on a molecular level, the structure of the reovirus attachment protein and its interaction with its host cell receptor.

We thank Dr Kenneth Tyler and Max Nibert for discussions and comments; Drs Edmund Choi and Jonathan Seidman for help with DNA sequencing and for discussions; Dr Stephen Harrison for discussion; Dr Angus Graham for continuous support; and Elaine Freimont for technical assistance. This work was supported by NIAID, PHS research grant 5 R01 AI13178, with partial support from the Shipley Institute of Medicine and the MRC of Canada (N.S. and Dr Angus Graham). R.B.-D. is supported by FCAC (Gouvernement du Québec).

Note added in proof: Since the submission of this manuscript, two other laboratories have published the sequence of the reovirus type 3 *S1* gene^{30,31}.

Received 17 September 1984; accepted 5 March 1985.

1. Smith, H., Skehel, J. J. & Turner, J. J. (eds) *The Molecular Basis of Microbial Pathogenicity*, 1-355 (Verlag-Chemie, Basel, 1980).
2. Fields, B. N. & Greene, M. I. *Nature* **300**, 19-23 (1982).
3. Both, G. W., Lavi, S. & Shatkin, A. J. *Cell* **4**, 173-180 (1975).
4. Smith, R. E., Zweerink, H. J. & Joklik, W. K. *Virology* **39**, 791-810 (1969).
5. Lee, P. W. K. *Double Stranded RNA Viruses* (eds Compans, R. H. & Bishop, D. H. L.) 193-206 (Elsevier, Amsterdam, 1983).
6. Kozak, M. J. *molec. Biol.* **156**, 807-820 (1982).
7. Cenatiempo, Y. et al. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1084-1088 (1984).
8. Ernst, H., & Shatkin, A. J. *Proc. natn. Acad. U.S.A.* **82**, 48-52.
9. Crick, F. H. C. *Acta crystallogr.* **6**, 689-697 (1953).
10. McLachlan, A. D. & Stewart, M. J. *molec. Biol.* **98**, 293-304 (1975).
11. Garnier, J., Osguthorpe, D. J. & Robson, B. J. *molec. Biol.* **120**, 97-120 (1978).
12. Ward, C. W. & Dopheide, T. A. *Aust. J. Biol. Sci.* **33**, 441-447 (1980).
13. Wilson, I. A., Skehel, J. J. & Wiley, D. C. *Nature* **289**, 366-373 (1981).
14. Joklik, W. K. *Reoviridae* (ed. Joklik, W. K.) 9-70 (Plenum, New York, 1983).
15. Spriggs, D. R. & Fields, B. N. *Nature* **297**, 68-70 (1982).
16. Spriggs, D. R., Bronson, R. T. & Fields, B. N. *Science* **220**, 505-507 (1983).
17. Goad, W. B. & Kanehisa, M. *Nucleic Acids Res.* **10**, 247-263 (1982).
18. Harington, W. F. & Rodgers, M. E. *A. Rev. Biochem.* **53**, 35-73 (1984).
19. McPhillips, T. H. & Ramig, R. F. *Virology* **135**, 428-439 (1984).
20. Hayes, E. C., Lee, P. W. K., Miller, S. & Joklik, W. K. *Virology* **108**, 147-155 (1981).
21. Luftig, R. B., Kilham, S., Hay, A. J., Zweerink, H. J. & Joklik, W. K. *Virology* **48**, 170-181 (1972).
22. Dayhoff, M. O. (ed.) *Atlas of Protein Sequence and Structure* Vol. 5 (National Biomedical Research Foundation, Washington, DC, 1978).
23. Sippel, A. E. *Eur. J. Biochem.* **37**, 31-40 (1973).
24. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
25. Deng, G. & Wu, R. *Nucleic Acids Res.* **9**, 4173-4188 (1981).
26. England, T. E., Bruce, A. G. & Uhlenbeck, O. C. *Meth. Enzym.* **65**, 65-74 (1980).
27. Messing, J. & Vieira, J. *Gene* **19**, 269-276 (1982).
28. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
29. Sarkar, G. et al. *J. Virol.* (in the press).
30. Nagata, L., Masri, S. A., Mah, D. C. W. & Lee, P. W. K. *Nucleic Acids Res.* **12**, 8699-8710 (1984).
31. Cashdollar, L. W., Chmelo, R. A., Wiener, J. R. & Joklik, W. K. *Proc. natn. Acad. Sci. U.S.A.* **82**, 24-28 (1985).

An immunoglobulin promoter displays cell-type specificity independently of the enhancer

John Foster, Jeannine Stafford & Cary Queen

Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205, USA

Recent studies in which cloned immunoglobulin genes were introduced into cultured cells have produced two significant findings. First, the genes are expressed after transfection into lymphoid cells but not non-lymphoid cells¹⁻³. Second, transcription of an immunoglobulin gene requires, in addition to the promoter region, an enhancer element located downstream of the transcription start site⁴⁻¹⁰. These findings raise the question of whether it is the promoter or the enhancer region that is responsible for the observed cell-type specificity. It has, in fact, been shown that immunoglobulin enhancers function only in lymphoid cells^{6,8,9}. We show here that the promoter for an immunoglobulin κ light-chain gene also is strongly specific for lymphoid cells. Our result re-emphasizes the importance of promoters relative to enhancers in determining which cells express which genes.

Because the κ light-chain enhancer is already known to be cell-type specific, any study of the specificity of the κ -chain promoter must separate the promoter from the enhancer. For this purpose, we constructed the plasmid pLX40 (Fig. 1a), which contains three important elements: (1) part of the MOPC 41 κ light-chain gene¹¹ extending from ~1.1 kilobases (kb) before the transcription initiation site to 2.1 kb after the initiation site; (2) the simian virus 40 (SV40) early promoter attached to the bacterial *neo* gene¹²; and (3) the polyoma virus large-T antigen gene and origin of replication. The part of the κ -chain gene used excludes the enhancer, which has been mapped⁵ to a region 1.6 kb further downstream. The κ -chain gene polyadenylation signal, also deleted in this construction, is replaced by an SV40 poly(A) signal to allow processing of κ -chain transcripts into stable messenger RNA^{4,5}. The role of the SV40 promoter is to provide an internal control for transcription from the κ -chain promoter. The polyoma segment allows the plasmid to replicate

to a high level in most mouse lymphoid and non-lymphoid cell lines (ref. 13 and our unpublished data), thus increasing the amount of RNA synthesized by increasing the amount of DNA template^{4,5,7}.

The κ -chain promoter cannot function without an enhancer, and so, to study transcription from the promoter, we inserted three different enhancers into pLX40 at an upstream *Eco*RI site. A 238-base pair (bp) fragment of polyoma (nucleotides 5,028-5,265)¹⁴ including the enhancer region¹⁵ was inserted in both orientations to form the plasmids pLX40P.1 and pLX40P.2. A 216-bp fragment of the Moloney murine sarcoma virus (MSV) long terminal repeat (nucleotides 210-426)¹⁶, containing the 73/72-bp repeat unit with full enhancer activity¹⁶, was inserted to form the plasmid pLX40M.1. For purposes of comparison, a 473-bp fragment of the κ -chain gene (nucleotides 3,689-4,161)¹⁷ containing the κ enhancer itself^{5,6} was also inserted in both orientations to form the plasmids pLX40K.1 and pLX40K.2. Although pLX40 already contains the SV40 enhancer as part of the SV40 origin (Fig. 1a), we inserted the enhancers from murine viruses to obtain maximal activation of the κ -chain promoter in mouse cells^{18,19}. In fact, pLX40 also contains the polyoma enhancer as part of the polyoma replicon, but the enhancer does not stimulate the κ -chain promoter in this configuration^{4,5}.

This set of six plasmids was transfected into antibody-secreting MPC 11 mouse myeloma cells and three lines of non-lymphoid mouse cells. Forty-eight hours after transfection, total RNA was extracted from the cells and assayed by the *S*₁ nuclease method^{20,21} in conditions of probe excess. For each RNA sample, one aliquot was tested with a 5' end-labelled, double-stranded probe specific for κ -chain transcripts and an equal aliquot with such a probe specific for *neo* transcripts (Fig. 1b). The two hybridizations were carried out at temperatures optimal for the respective probes, and the two *S*₁ nuclease-treated reactions combined before loading on a gel.

Figure 2 shows the results of a typical *S*₁ experiment after transfection of the plasmids into MPC 11 cells. Bands of the intact *neo* and κ probes were present (Fig. 2), because of some re-annealing of the probes during the hybridizations and to some upstream transcription from pBR322 sequences (data not shown). In addition, a strong band was present beneath the *neo* probe. The size of this fragment corresponds to transcripts

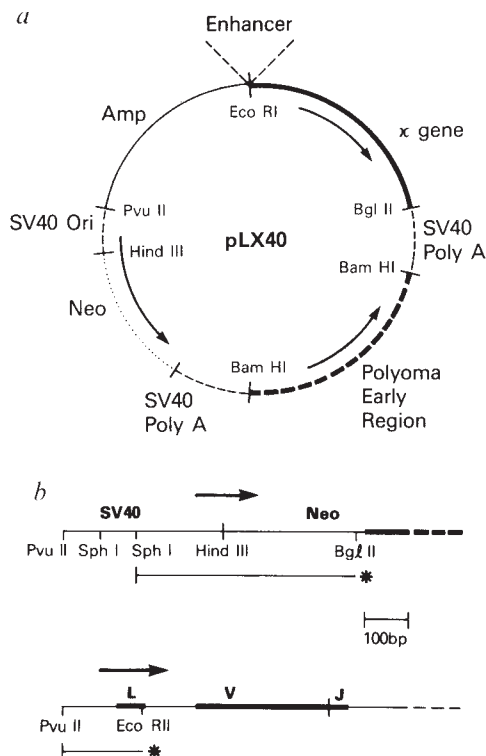


Fig. 1 **a**, Schematic diagram of the plasmid pLX40. Sources of DNA fragments are indicated as follows: pBR322, solid line; SV40, dashed line; *neo* gene, dotted line; polyomavirus, heavy dashed line; κ gene, heavy line. The structure of the pBR322-*neo*-SV40 segment is as described elsewhere¹²; the κ segment extends from the upstream *Eco*RI site to the *Bgl*II site in the intron^{11,17}; the SV40 segment after the κ -chain gene is the same as the SV40 segment after the *neo* gene, and the polyoma segment is the same as used previously⁴. Arrows indicate the approximate initiation sites and directions of transcription for the SV40 early promoter, the polyoma early promoter and the κ promoter. Enhancers were inserted at the *Eco*RI site by means of linkers to form other plasmids. **b**, Strategy for S₁ nuclease assays. The *neo*-specific and κ -specific probes are shown underneath the respective genes, with the ³²P-labelled ends denoted by asterisks. Initiation sites and direction of SV40 early and κ transcription are shown by arrows, and translated regions of DNA are drawn in heavy line. L, leader segment; V, variable region; J, joining segment. The plasmid pLX40 and derivatives were constructed by standard methods; details are available on request to the authors. The *neo* and κ probes were prepared as described previously² with specific activities of $\sim 2 \times 10^6$ d.p.m. per pmol fragment.

initiating from the SV40 promoter; the band runs as a doublet because of heterogeneity in the SV40 early transcription start site²². Another strongly protected fragment was present in all the samples below the κ probe. This fragment is characteristic of transcripts initiated by the κ -chain promoter 25 bp before the first codon of the gene^{4,7}. Neither the *neo*-specific nor κ -specific fragment were protected by RNA extracted from untransfected cells². The transcription from the κ -chain promoter on pLX40 presumably resulted from the effect of the SV40 enhancer. κ -chain transcription was stimulated $\sim$ fivefold more by any of the murine enhancers (Fig. 2).

The set of plasmids was also transfected into a transformed line of liver cells BNL 1ME A.7R.1 (ATTC TIB 75) as a representative non-lymphoid cell line. Strong *neo*-specific bands were detected after S₁ analysis, showing that the plasmids were successfully transfected into the liver cells and could be transcribed. Indeed, more *neo* RNA was produced in the liver cells than in the myeloma cells, probably because of a higher transfection efficiency. Nonetheless, κ -chain transcripts were almost undetectable in the liver cell RNAs, although longer photographic exposures did reveal a small amount of κ -chain specific transcript from pLX40P.1, P.2 and M.1. Very similar results

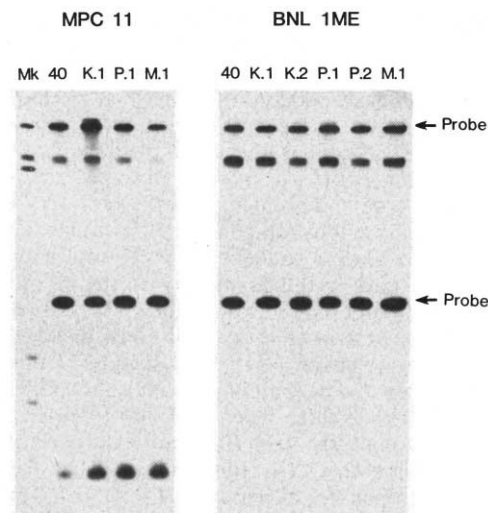


Fig. 2 S₁ nuclease assay of RNA from transfected cells. Cells of strains MPC 11 or BNL 1ME A.7R.1 were transfected with DNA of the indicated plasmids. Total RNA was extracted from the cells after 48 h and an equal aliquot of each RNA sample was hybridized at 53 °C to a *neo*-specific probe and at 48 °C to a κ -specific probe. After S₁ digestion, the reactions were combined and analysed on a polyacrylamide gel. The upper and lower bands marked probe co-migrated respectively with undigested *neo* and κ probe. The markers are from a labelled, denatured *Hinf*I digest of the plasmid pSV2-*neo*¹² and have lengths in nucleotides of 517, 396, 362, 134, 109 and 83. Mk, markers; 40, plasmid pLX40; K.1, pLX40K.1; K.2, pLX40K.2; P.1, pLX40P.1; P.2, pLX40P.2; M.1, pLX40M.1. Transfection of cells, extraction of RNA and S₁ assays were performed as described previously².

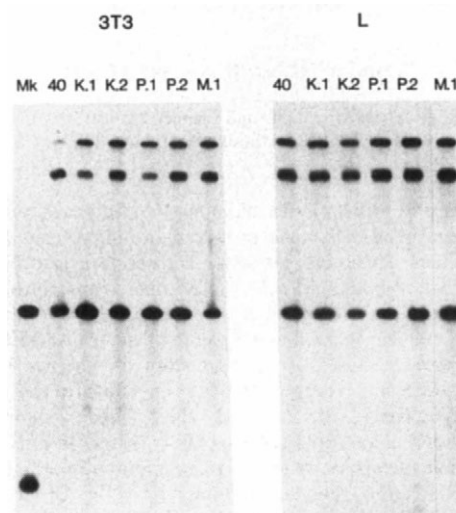


Fig. 3 S₁ nuclease assay of RNA from transfected cells. The assay was performed as in Fig. 2. Lane Mk is an S₁ analysis with the κ probe of RNA from MPC 11 cells transfected with pLX40P.1, performed in parallel with the other reactions. The cells were transfected as before², with the addition of a dimethyl sulphoxide shock step as described elsewhere²⁶.

were obtained when the plasmids were transfected into mouse L and 3T3 cells (Fig. 3): the SV40 promoter transcribed the *neo* gene strongly on all the plasmids, but κ -chain transcription was very weak from pLX40P.1, P.2 and M.1 and undetectable from the other plasmids. The MSV enhancer is known to function in both L and 3T3 cells^{16,18} and the polyoma enhancer functions in a wide range of cells, including 3T6 cells¹⁹, and apparently in 3T3 cells where polyomavirus can replicate¹⁴. Hence, the very low level of κ -chain transcripts detected from pLX40P.1, P.2 and M.1 suggests that it is the κ -chain promoter itself which is intrinsically inactive in the non-lymphoid cells.

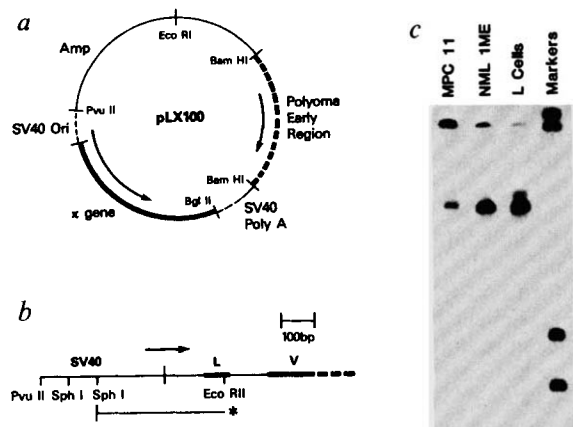


Fig. 4 *a*, Schematic diagram of the plasmid pLX100. Transcripts and sources of the DNA fragments are indicated as in Fig. 1. To construct the plasmid, a fragment of the plasmid pLX40 extending from the *Pvu*II site upstream of the κ gene (Fig. 1*b*) to the *Bam*HI site after the gene was inserted between the unique *Hind*III and *Bam*HI sites of the plasmid pSV2-neo¹²; the *Bam*HI polyoma fragment of pLX40 was then inserted at the *Bam*HI site of the new plasmid. *b*, Probe used for *S*₁ nuclease assays of RNA from pLX100. The probe is shown beneath the region of pLX100 where the SV40 origin and the κ gene are joined, with the ³²P-labelled end denoted by an asterisk. The arrow indicates transcription from the SV40 early promoter. *c*, *S*₁ nuclease assay of RNA from transfected cells. Cells of the indicated strains were transfected with pLX100 DNA, and extracted RNA was analysed at 53 °C with the probe shown in *b*. The topmost band in each lane co-migrated with undigested probe. The markers are part of a *Hinf*I digest of pSV2-neo DNA and have sizes in nucleotides of 396, 362, 134 and 109.

However, an alternative explanation is that κ -chain transcripts are synthesized at a higher level but are unstable in non-lymphoid cells. To eliminate this possibility, we constructed the plasmid pLX100 (Fig. 4*a*) in which the SV40 early promoter was attached to the same deleted κ -chain gene used in the pLX40 plasmids. The new plasmid was transfected into MPC 11, BNL 1ME and L cells, and extracted RNA was assayed by the *S*₁ method, using a suitable probe (Fig. 4*b*). In contrast to transcripts of the κ -chain gene initiating from its own promoter on the pLX40 plasmids (Figs 2, 3), transcripts of the κ -chain gene initiating from the SV40 promoter gave a strong signal (Fig. 4*c*). In fact, these transcripts were detected at a higher level in the non-lymphoid cells than in the myeloma cells, probably because of the higher transfection efficiency of the non-lymphoid cells seen above. In any case, the results show clearly that transcripts of this κ -chain gene are stable in non-lymphoid cells, and we therefore conclude that the κ -chain promoter initiates transcription very weakly in those cells.

We also transfected the pLX40 plasmids into lines of pre-B, B and T cells. We found that both the promoter and enhancer function in these cells, but not as efficiently as in myeloma cells, the activities in T cells being especially weak (manuscript in preparation).

The precise sequences responsible for the observed cell-type specificity of the κ -chain promoter may *a priori* be at a considerable distance from the actual transcription start site. However, we have shown that sequences more than 100 bp upstream of the start site are not required for high-level transcription in myeloma cells (C. Stauber and C.Q., unpublished data), and sequences downstream of the variable segment are also not required⁷. Hence, it seems likely that the actual sequences inducing high-level transcription specifically in B-lymphoid cells are fairly close to the transcription start site. Excellent candidates for these sequences would be two short regions of homology found in all light-chain genes examined, about 50–100 bp upstream of the start site^{23,24}. Similar conclusions have recently been reached by others²⁵.

Received 29 November 1984; accepted 12 March 1985.

- Falkner, F. & Zachau, H. *Nature* **298**, 286–288 (1982).
- Stafford, J. & Queen, C. *Nature* **306**, 77–79 (1983).
- Gillies, S. & Tonegawa, S. *Nucleic Acids Res.* **11**, 7981–7997 (1983).
- Queen, C. & Baltimore, D. *Cell* **33**, 741–748 (1983).
- Queen, C. & Stafford, J. *Molec. cell. Biol.* **4**, 1042–1049 (1984).
- Picard, D. & Shaffner, W. *Nature* **307**, 80–83 (1984).
- Bergman, Y., Rice, D., Grosschedl, R. & Baltimore, D. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7041–7045.
- Gillies, S. D., Morrison, S. L., Oi, V. T. & Tonegawa, S. *Cell* **33**, 717–728 (1983).
- Banerji, J., Olson, L. & Schaffner, W. *Cell* **33**, 729–740 (1983).
- Neuberger, M. *EMBO J.* **2**, 1373–1379 (1983).
- Seidman, J. G. & Leder, P. *Nature* **276**, 79–795 (1978).
- Southern, P. J. & Berg, P. *J. molec. appl. Genet.* **1**, 327–341 (1982).
- Deans, R., Denis, K. A., Taylor, A. & Wall, R. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1292–1296 (1984).
- Toozee, J. (ed.) *DNA Tumor Viruses* (Cold Spring Harbor Laboratory, New York, 1980).
- de Villiers, J. & Schaffner, W. *Nucleic Acids Res.* **9**, 6251–6264 (1981).
- Laimins, L. A., Gruss, P., Pozzatti, R. & Khoury, G. *J. Virol.* **49**, 183–189 (1984).
- Max, E. E., Maizel, J. V. & Leder, P. *J. biol. Chem.* **256**, 5116–5120 (1981).
- Laimins, L. A., Khoury, G., Gorman, C., Howard, B. & Gruss, P. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6453–6457 (1982).
- de Villiers, J., Olson, L., Tyndall, C. & Schaffner, W. *Nucleic Acids Res.* **10**, 7965–7975 (1982).
- Berk, A. J. & Sharp, P. A. *Cell* **12**, 721–732 (1977).
- Weaver, R. F. & Weissmann, C. *Nucleic Acids Res.* **6**, 1175–1192 (1979).
- Reddy, V. B., Ghosh, P. K., Lebowitz, P., Piatak, M. & Weissman, S. M. *J. Virol.* **30**, 279–296 (1979).
- Falkner, F. G. & Zachau, H. G. *Nature* **310**, 71–74 (1984).
- Parslow, T. G., Blair, D. L., Murphy, W. J. & Granner, D. K. *Proc. natn. Acad. Sci. U.S.A.* **81**, 2650–2654.
- Falkner, F. G., Neumann, E. & Zachau, H. *Hoppe-Seyler's Z. physiol. Chem.* **365**, 1331–1343.
- Lopata, M. A., Cleveland, D. W. & Sollner-Webb, B. *Nucleic Acids Res.* **12**, 5707–5717 (1984).

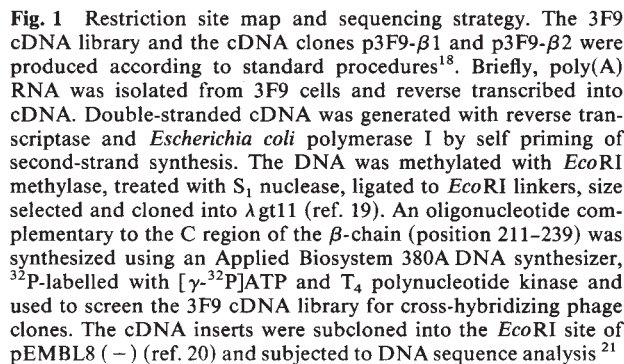
Identical V_{β} T-cell receptor genes used in alloreactive cytotoxic and antigen plus I-A specific helper T cells

Fabio Rupp, Hans Acha-Orbea, Hans Hengartner, Rolf Zinkernagel & Rolf Joho

Abteilung für Experimentelle Pathologie, Universitätsspital, Sternwartstrasse 2, 8091 Zürich, Switzerland

T lymphocytes involved in the cellular immune response carry cell-surface receptors responsible for antigen and self recognition. This T-cell receptor molecule is a heterodimeric protein consisting of disulphide-linked α - and β -chains with variable (V) and constant (C) regions^{1–5}. Several complementary DNA and genomic DNA clones have been isolated and characterized^{6–12}. These analyses showed that the genomic arrangement and rearrangement of T-cell receptor genes using V_T , diversity (D_T), joining (J_T) and C_T gene segments is very similar to the structure of the known immunoglobulin genes (see ref. 13 for review). We have isolated two cDNA clones from an allospecific cytotoxic T cell, one of which shows a productive V_{β} - J_{β} - $C_{\beta}1$ rearrangement without an intervening D_{β} segment. This V_{β} gene segment is identical to the V_{β} gene expressed in a helper T-cell clone specific for chicken red blood cells and H-2I^b (ref. 14). The other clone carries the $C_{\beta}2$ gene of the T-cell receptor, but the $C_{\beta}2$ sequence is preceded by a DNA sequence that does not show any similarity to V_{β} or J_{β} sequences.

An allospecific cytotoxic T-cell clone 3F9 of BALB/c origin with specificity for H-2D^b has been established¹⁵. To isolate cDNA clones bearing sequences for the different chains (α and β) of the T-cell receptor molecule(s), we constructed a cDNA library in the bacteriophage λ gt11. A synthetic oligonucleotide complementary to part of the C region of the β -chain was synthesized and used to screen the λ gt11-3F9 cDNA library for cross-hybridizing cDNA clones. Twenty such clones were isolated from ~100,000 recombinant phages. Southern blot analysis of cloned phage DNA indicated two groups of cDNA clones, probably corresponding to $C_{\beta}1$ and $C_{\beta}2$ sequences. One clone bearing the longest DNA insert of each group was chosen, and the cDNA was subcloned into the *Eco*RI site of pEMBL8 (–). The restriction map and sequencing strategy of the two clones is outlined in Fig. 1. Sequence analysis of p3F9- β 1 confirmed that this clone was derived from a $C_{\beta}1$ -bearing transcript



To confirm the absence of a D gene segment in 3F9, we determined the 3' end of the V_β gene by comparing it with a germline V_β gene. A genomic unrearranged V_β gene had previously been cloned from B10.A liver DNA⁸. This germline V_β gene is homologous at its 3' end with the V_β gene found in 3F9 and LB2 (Fig. 2). The last 17 nucleotides of the 2B4.71 germline V-region gene are identical to the sequence immediately preceding the $J_{\alpha 1.1}$ gene segment in p3F9- $\beta 1$. The 2B4.71 coding

Figure 3 shows the sequence comparison of the 5' end of p3F9- $\beta 2$ and the genomic DNA sequence found 50-90 nucleotides upstream of the $D_{\beta 2.1}$ gene segment. The end of this homologous stretch of DNA is characterized by a consensus sequence of a donor site for RNA splicing. The p3F9- $\beta 2$ cDNA is derived from a mRNA that may have arisen by splicing a RNA transcript from this genomic DNA sequence directly to the $C_{\beta 2}$ coding sequence. This means that the normal donor sites for RNA splicing which are present at the 3' ends of the various $J_{\beta 2}$ gene segments have been omitted. The reasons for this aberrant RNA splicing are not clear; however, Southern blot analysis of genomic 3F9 DNA indicates that DNA deletions have occurred near the $J_{\beta 2}$ region (M. Schilham, R. Lang and

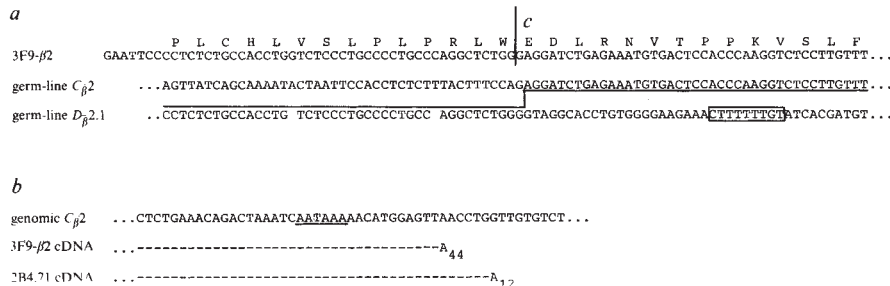
P K F L I L I G Q E G Q K L T L K K C Q Q N F N H D T M Y W Y R Q D
 3F9-81 GAATTCGCCCAAATTCCTGATTGGTCAGGAAGGGCAAAACTGACCTTGAATGTCAACAGAAATTCATCATGATACAATGTACTGGTACCGACAGGAT
 LB2 -----

 S G K G L R L I Y Y S I T E N D L Q K G D L S E G Y D A S R E K K
 3F9-81 TCAGGGAAGAGGATTGAGACTGATCTACTATTCAATAACTGAAACGACTCTTCAAAAAGCGATCTATCTGAAAGGCTATGATGCGTCTCGAGAGAAGAAGT
 LB2 -----

 S S F S L T V T S A Q K N E M A V F L C A S S L N T E V F F G
 3F9-81 CATCTTTTTCCTCACTGTGACATCTGCCCAAGAAGCAGAGTGGCGTTTCTCTGTCCACGACTCTA AACACAGAGTCTCTTTTGG
 LB2 -----A-----A-AGACTGGCG-CTG-----ACGC-G-A-TT

 germ-line 2B4.71 ...-AC-----CTCCACAGCATTGAATGTGCATCTCTCC

Fig. 3 Sequence of p3F9- β_2 . *a*, The 5' terminal sequence and the beginning of the $C_{\beta 2}$ sequence are shown and compared with the germline sequence^{10,11} of $C_{\beta 2}$ and a DNA sequence 5' to the $D_{\beta 2.1}$ gene segment¹². The conserved nonamer preceding the $D_{\beta 2.1}$ gene segment is boxed. The sequence found in the cDNA plasmid p3F9- β_2 is indicated by a line. *b*, The 3' terminal sequence of two cDNAs and the corresponding genomic sequence is shown. The polyadenylation signal ATAAA is underlined.



R.J., unpublished observations). Therefore, it is possible that the $J_{\beta 2}$ gene cluster has been deleted and a RNA transcript starting upstream of this deletion is now differently spliced. This finding implies that a promoter activity should be present 5' of the J_{β} gene region. Similar cDNA molecules have been isolated from other T cells, but the cDNA molecules from other T cells contain J_{β} sequences 5' of their respective C_{β} sequences^{12,16}.

p3F9- β_2 contains a long tail of A residues at its 3' end. Therefore, we could determine the position of poly(A) addition in the primary transcript. Figure 3*b* compares the 3' ends of two different cDNA molecules both derived from a β_2 mRNA with the corresponding genomic $C_{\beta 2}$ gene^{10,11}. Although both transcripts are identical in the heteropolymeric region of the 3' end, the site where polyadenylation begins varies by four to six nucleotides. If this does not result from allelic differences in the 3' untranslated regions of 3F9 (BALB/c) and 2B4.71 (B10.A), then this comparison indicates that the exact position of polyadenylation is not determined on different transcripts even of the same gene.

We thank Dr Martin Billeter for helpful discussions and comments, Jolanda Brecher for technical assistance and Béatrice Borter and Rosina Caprez for preparation of the manuscript.

This work was supported by grant 3.765-0.80 of the Schweizerischer Nationalfond to R.J. and grant 3.329-0.82 to H.H.

Received 26 February; accepted 24 April 1985.

- Acuto, O. *et al. Cell* **34**, 717-726 (1983).
- Kappler, J., Kubo, R., Haskins, K., White, J. & Marrack, P. *Cell* **34**, 727-737 (1983).
- McIntyre, B. W. & Allison, J. P. *Cell* **34**, 739-746 (1983).
- Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
- Méuer, S. C. *et al. J. exp. Med.* **157**, 705-719 (1983).
- Hedrick, S. M., Cohen, D. I., Nielsen, E. A. & Davis, M. M. *Nature* **308**, 149-158 (1984).
- Yanagi, Y. *et al. Nature* **308**, 145-149 (1984).
- Chien, Y., Gascoigne, N. R. J., Kavalier, J., Lee, N. E. & Davis, M. M. *Nature* **309**, 322-326 (1984).
- Siu, G. *et al. Cell* **37**, 393-401 (1984).
- Malissen, M. *et al. Cell* **37**, 1101-1110 (1984).
- Gascoigne, N. R. J., Chien, Y., Becker, D. M., Kavalier, J. & Davis, M. M. *Nature* **310**, 387-391 (1984).
- Siu, G. *et al. Nature* **311**, 344-350 (1984).
- Joho, R., Nottenberg, C., Coffman, R. L. & Weissman, I. L. *Curr. Topics dev. Biol.* **18**, 15-58 (1983).
- Patten, P. *et al. Nature* **312**, 40-46 (1984).
- Acha-Orbea, H., Zinkernagel, R. M. & Hengartner, M. *Eur. J. Immun.* **15**, 31-36 (1985).
- Yoshikai, Y. *et al. Nature* **312**, 521-524 (1984).
- Kranz, D. M. *et al. Nature* **313**, 752-755 (1985).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Young, R. A. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **80**, 1194-1198 (1983).
- Dente, L., Cesareni, G. & Cortese, R. *Nucleic Acids Res.* **11**, 1645-1653 (1983).
- Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499-560 (1980).

Detection and sequence of mutations in the factor VIII gene of haemophiliacs

Jane Gitschier, William I. Wood,
Edward G. D. Tuddenham*, Marc A. Shuman†,
Therese M. Goralka, Ellison Y. Chen
& Richard M. Lawn

Genentech, Inc., 460 Point San Bruno Boulevard,
South San Francisco, California 94080, USA

* Haemophilia Centre, Academic Department of Haematology,
Royal Free Hospital School of Medicine, London NW3 2QG, UK

† Hematology Section, Cancer Research Institute, University of
California, San Francisco, California 94143, USA

The most common inherited bleeding disorder in man, haemophilia A, is caused by defect in factor VIII, a component in the blood coagulation pathway. The X-chromosome-linked disease almost certainly stems from a heterogeneous collection of genetic lesions. Because, without proper treatment, haemophilia can be a fatal disease, new mutations are necessary to account for its constant frequency in the population¹. In addition, haemophilia A displays a wide range of severity, and some 15% of haemophiliacs generate high levels of antibodies against factor VIII ('inhibitor patients'). The present work elucidates the molecular genetic basis of haemophilia in some individuals. Using the recently cloned factor VIII gene as a probe^{2,3}, we have identified two different nonsense point mutations in the factor VIII gene of haemophiliacs, as well as two different partial deletions of the gene. Our survey of 92 haemophiliacs indicates no firm correlation between antibody (inhibitor) production and gross gene defects.

In an initial study involving 36 haemophilia patients, DNA samples were digested and analysed by Southern blot hybridization to the cloned factor VIII gene. Each DNA sample was digested with 4-11 restriction enzymes to search for variant hybridization patterns. One result of this endeavour has been the discovery of a common restriction fragment length polymorphism with *Bcl*I which allows the tracking of specific X chromosomes in many families. The use of this factor VIII DNA polymorphism in gene localization and prenatal diagnosis is reported elsewhere⁴. The restriction enzyme *Taq*I proved particularly useful for revealing differences in the DNA of individual haemophiliacs. This result is not surprising, because the recognition sequence for *Taq*I contains the CpG dimer, which is the major site of methylation of human DNA. CpG dimers are probably mutation hot spots because of subsequent deamination of methylcytosine to thymidine, resulting in C to T transitions⁵. The location of the affected restriction enzyme sites and the sequence of the factor VIII gene² alerted us to the possibility of in-frame stop codon mutations generated by C to T transitions in the *Taq*I sequence. Five of the seven *Taq*I sites in exons of the factor VIII gene can be converted to nonsense mutations by such a mechanism. We therefore concentrated on the screening of 92 haemophilic DNA samples with *Taq*I. (The restriction enzyme *Msp*I also contains CpG in its recognition sequence, but no variant hybridizing bands were found in 20 samples.) Case-specific *Taq*I site mutations were discovered in the DNA of four severe haemophiliacs (none of these four patients has any detectable factor VIII activity).

Figure 1 shows a representative Southern blot of 11 DNA samples digested with *Taq*I and probed with the 9-kilobase (kb) cloned factor VIII complementary DNA. Distinct variant fragments can be seen in haemophilic patients H2 and H6. The *Taq*I site change in patient H2 was mapped to exon 24 of the 26 exons comprising the human factor VIII gene³. A hybridiz-

ation probe containing only exons 21–26 simplifies the band pattern and was subsequently used in the analysis of DNA from haemophilic H2 and members of his family (Fig. 2a). Absence of the *TaqI* site in exon 24 produces a 4.2-kb hybridizing band found only in patient H2. His mother as well as his father, sister and two brothers display the normal 1.4- and 2.8-kb fragments, but not the 4.2-kb fragment, suggesting that this new mutation occurred in the X chromosome of individual H2, a result consistent with the lack of previous haemophilia in this family.

A region of H2 DNA which includes this mutation was cloned into bacteriophage λ , and exon 24 was sequenced. The DNA sequence was identical to that of normal DNA analysed previously², except that the *TaqI* site had changed from TCGA to TTGA in the H2 DNA. This mutation generates an in-frame stop codon in the factor VIII gene 124 amino acids before the normal chain termination. Because this mutation was found in only 1 case out of 92 haemophilic and 80 non-haemophilic X chromosomes examined, and because it first appeared in the only affected member of this family, we propose that this non-sense mutation is likely to be the cause of haemophilia in this individual. In addition to having a new haemophilia mutation,

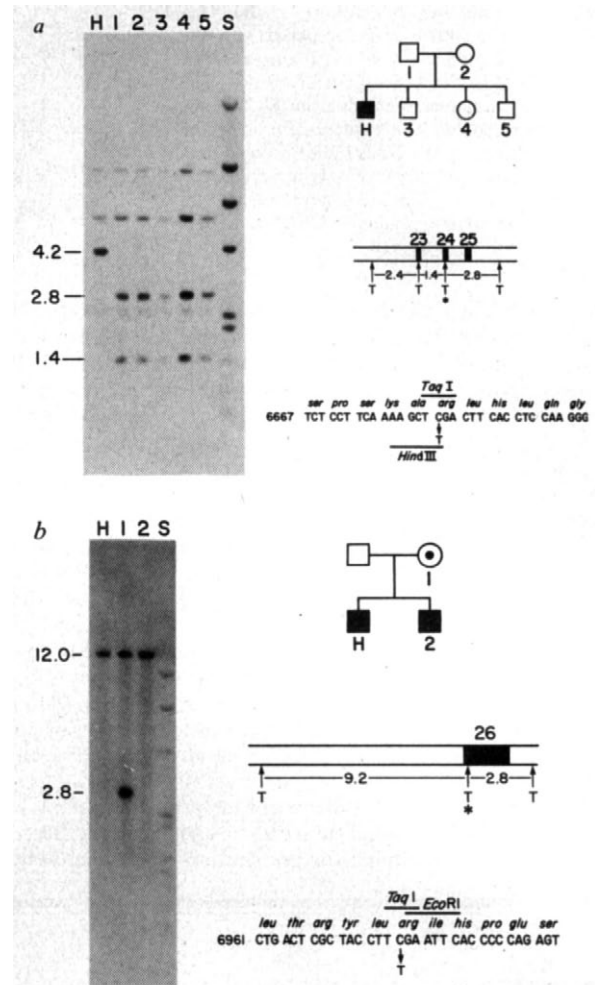


Fig. 2 Analysis of the factor VIII gene of patient H2 and patient H22. **a**, A Southern blot of DNA from patient H2 and his family members was probed with a cDNA 783-bp *ApaI*/*EcoRI* fragment that includes all of exons 22–25 and parts of exons 21 and 26. Lanes H (haemophilic), 1, 2, 3, 4 and 5 correspond to family members in the accompanying pedigree [□, normal male; ■, haemophilic male; ○, normal female; ⊙, carrier female (not present on this particular pedigree)]. The diagram below the pedigree indicates with an asterisk the location of the missing *TaqI* site in exon 24. Below this is shown the sequence of normal DNA² and the changed base in H2 DNA that results in the loss of a *TaqI* site, the generation of a *HindIII* site (verified by a Southern blot, data not shown), and the occurrence of an in-frame stop codon TGA. The nucleotide number to the left follows ref. 2. **b**, A Southern blot of *TaqI*-digested DNA from patient H22, his mother and brother was probed with a 1.9-kb *EcoRI* cDNA fragment (from clone λ 10.3) that includes almost all of exon 26 (ref. 2). The mother (lane 1) has the normal 2.8-kb *TaqI* allele as well as the variant 12.0-kb allele, which she shares with her two haemophilic sons. The *TaqI* site that is missing in H22 DNA, and the normal DNA sequence and the altered nucleotide in H22 are indicated.

Methods. Blot hybridization was performed as described in Fig. 1 legend. For sequence analysis, DNA from patient H2 was preparatively digested with *Bam*HI. A 15-kb region, known to contain exon 24, was gel isolated and cloned into bacteriophage vector Charon 30 (ref. 11). An exon 24-containing 600-bp *Hae*III fragment of the resulting λ clone was subcloned into M13 and sequenced by dideoxy-chain termination procedures. The C to T transition is the only difference found between the normal and H2 sequence. Similarly, for patient H22, a 5.3-kb *Bam*HI fragment was cloned into vector Charon 30. A 1.6-kb *Sst*I/*Bam*HI fragment from this cloned DNA was then subcloned into M13 for sequencing from the *Sst*I site for ~600 bp to read through the *TaqI* and *Eco*RI sites.

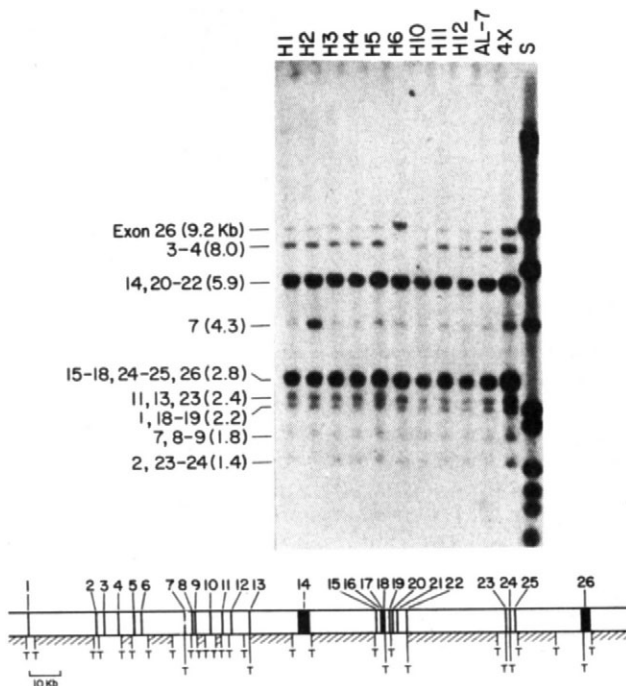


Fig. 1 Southern blot of haemophilia DNA samples showing two variations with the enzyme *TaqI*. The blot was hybridized to the full-length factor VIII probe (9.0-kb *Clal*/*Sst*II fragment of pSAT.8cl; ref. 2) which contains all coding sequences and the 3'-untranslated region. Each hybridizing band corresponds to at least one exon; band sizes in kb are shown in parentheses. The lanes containing the haemophilic samples are indicated by a number preceded by the letter H. The control lanes labelled AL-7 and 4X contain DNA prepared from cell lines which served as a source of the factor VIII mRNA and factor VIII genomic clones, respectively^{2,3}. Lane S contains ³²P-end-labelled *Hind*III-digested λ DNA and *Hae*III-digested Φ X174 DNA as standards. The diagram shows the location of *TaqI* sites (T) in exons (indicated by a lower T) and introns (indicated by hatchmarks). The intron areas indicated by hatchmarks are not mapped with respect to *TaqI* sites as they are not detected by the factor VIII cDNA probe. Bands predicted to correspond to exons 5–6 and 12 were not detected in the blot shown; the size of the fragment corresponding to exon 10 is uncertain.

Methods. Genomic DNA (5 μ g), prepared either from peripheral blood lymphocytes of haemophilics or from cell lines, was electrophoresed and blotted onto a nitrocellulose filter. The filter was hybridized to calf thymus primed ³²P-labelled probe in 50% formamide at 42 °C overnight and washed at 60 °C in 0.2XSSC before autoradiography^{11,12}.

H2 is also a significant case because the patient, a severe haemophilic with <1% factor VIII activity, has developed an extremely high titre of anti-factor VIII antibody (8,000 Bethesda units).

A different premature termination mutation was found in patient H22. The *TaqI* site normally found in exon 26 has been lost, leading to the replacement of the normal 2.8-kb hybridizing band with a 12-kb band (Fig. 2b). Haemophilic H22 and his haemophilic brother (Fig. 2b, lanes H and 2, respectively) display the 12-kb fragment, while DNA from their mother (lane 1), who is a carrier of haemophilia, contains the affected X chromosome with the 12-kb fragment, and a normal X chromosome with the 2.8-kb fragment. Cloning of H22 DNA and sequence analysis demonstrated a TCGA to TTGA transition of the *TaqI* site in exon 26, again generating an in-frame stop codon which we propose to be the basis of this haemophilia. Although this sequence predicts the synthesis of a protein only 26 amino acids shorter than normal factor VIII, such a truncated protein may well be inactive. The absent portion of the molecule is contained in one of the active subunits (relative molecular mass 80,000)^{2,6,7} of the mature protein and includes a cysteine residue which is conserved in the duplicated C domains^{3,7} and which may be important for the stability or active conformation of the protein. (Although patient H22 is a severe haemophilic, he does not produce antibodies against factor VIII.)

In the case of patient H6, the missing *TaqI* site is located within an intron, about 400 base pairs (bp) 3' to exon 2 and 2,000 bp 5' to exon 3. Figure 3a illustrates the inheritance of this mutation in the H6 family. The normal *TaqI* fragments of 1.5 and 8.0 kb detected using a subgenic probe are replaced by a 9.5-kb fragment in H6 DNA. The haemophilic's mother (Fig. 3a, lane 1) is an obligate carrier of haemophilia because her father (and four uncles) was a haemophilic. The carrier status of the haemophilic's sister and her two daughters (Fig. 3a, lanes 4-6) was unknown before this experiment. The results from this Southern blot indicate that the sister and one niece are carriers of haemophilia, whereas the other niece is not. We believe this to be the first carrier detection of haemophilia A made by DNA diagnosis using a probe within the factor VIII gene. The H6 DNA surrounding the missing *TaqI* site was cloned and sequenced and, again, a TCGA to TTGA transition in the *TaqI* site was found. This mutation does not create a sequence resembling consensus splice donor or acceptor sites^{3,8}, and is not located near the intron-exon borders. Thus, it is not apparent whether this sequence change leads to aberrant RNA splicing, and therefore whether it is directly responsible for the disease in this family.

The final variant involves a new *TaqI* site in haemophilic H40 DNA. The site of mutation is in the last intron of the factor VIII gene, ~20 kb 3' of exon 25 and 1.7 kb 5' of exon 26. A TCAA to TCGA change creates the new *TaqI* site. Sequence analysis of this intron also revealed a second difference when compared with the sequence of our previously isolated genomic clones³. A T to C transition appears 22 nucleotides 5' to the *TaqI* site mutation. We do not know whether this particular T-C variation is common among individuals. However, both of these sequence changes occur in a region homologous to the *Alu* family of interspersed repetitive DNA⁹. Furthermore, there is no obvious connection between these sequence anomalies in the H40 family and the disease; no consensus splice sites are created by these changes.

To summarize the *TaqI* survey, we have identified four independent *TaqI* site mutations in the factor VIII genes of haemophiliacs. Three of these involve the loss of a *TaqI* site (two of these changes are in exons) and one involves the introduction of a *TaqI* site in an intron. In addition to the 92 haemophilic DNA samples examined with *TaqI*, 80 normal (non-haemophilic) chromosomes have been screened by hybridization of Southern blots of *TaqI*-digested DNA with a full-length factor VIII probe. Of those 172 chromosomes, only the 4 *TaqI* site mutations have been observed, suggesting that each one is a rare variant ($\leq 0.6\%$).

We propose that the in-frame stop codons found in exon 24 of the factor VIII gene in patient H2 and in exon 26 in H22 are the cause of haemophilia A in these families. Additional mutations cannot be ruled out, however. The phenotype could be due to inactivity or instability of the truncated protein or an effect of premature termination on the stability of the messenger

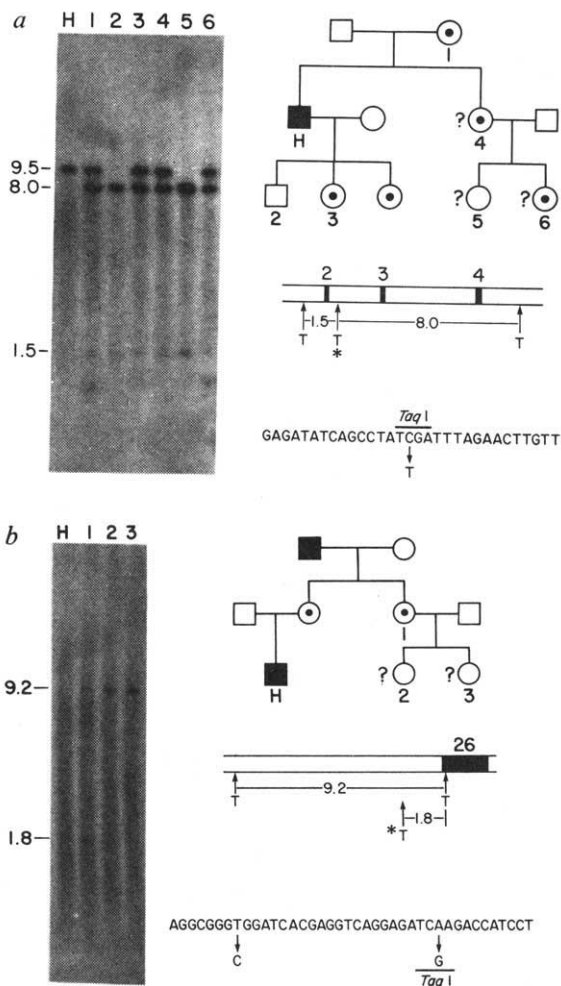


Fig. 3 Analysis of the factor VIII gene of patients H6 DNA and H40. **a**, A Southern blot of *TaqI*-digested DNAs from patient H6 (H) and his family members was probed with 347-bp *EcoRI* cDNA fragment that includes only exons 2, 3 and 4 (ref. 2). The question marks on the pedigree indicate that before this experiment the carrier status of 4, 5 and 6 was unknown. The diagram below the pedigree shows the location of the *TaqI* site (T*) 3' of exon 2 that is absent in the DNA of patient H6. Also shown is the sequence of normal DNA and the C to T change in the sequenced H6 DNA. **b**, A Southern blot of *TaqI*-digested DNA from patient H40, his aunt and two cousins was hybridized to a ³²P-labelled 500-bp *SstI/EcoRI* genomic fragment prepared from cosmid clone p613 (ref. 3). The unusual *TaqI* site in H40 DNA was mapped to an intron 1.7 kb 5' of exon 26, as shown by T* in the schematic diagram. The 1.8-kb *TaqI* fragment is present in DNA samples from the haemophilic and the aunt, who also has one normal allele, but is not found in DNA from the female cousins.

Methods. From the H6 DNA, the appropriate 7.0-kb *BclI* fragment was prepared and cloned into *BamHI* arms of the Charon 30 vector. A 600-bp *EcoRI/KpnI* fragment from this H6 λ clone, as well as the corresponding normal fragment from cosmid clone p624 (ref. 3), were subcloned into M13 and sequenced by dideoxy-chain termination methods. The C to T change in the *TaqI* site was the only difference in sequence between these two samples. From patient H40, *BamHI*-digested DNA was preparatively run on a gel, and a 5.3-kb region, known to contain the new *TaqI* site, was isolated and cloned into Charon 30 vector. From this λ clone, a 1.0-kb *SstI/SmaI* fragment was isolated, and along with the analogous normal fragment, a 1.0-kb *SstI/SmaI* fragment from cosmid clone p613 (ref. 3), was cloned into M13 and sequenced.

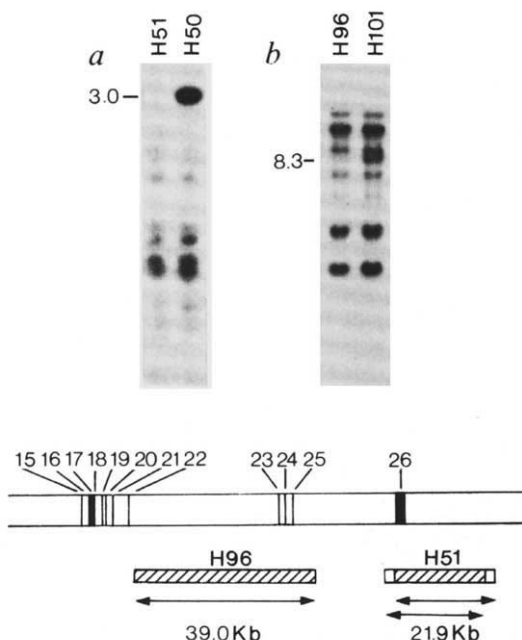


Fig. 4 Map of the deletions in the factor VIII genes of patients H51 and H96. **a**, A Southern blot of *RsaI*-digested genomic DNA from a number of haemophiliacs, including H51, was hybridized with the full-length factor VIII probe. A 3.0-kb *RsaI* fragment present in all other haemophilic samples tested is absent in H51 DNA (only one non-deletion DNA, H50, is shown). To map the H51 deletion, three single-copy genomic fragments were prepared from a cosmid clone p625, which extends 37 kb beyond the 3' end of the factor VIII gene. In addition, two genomic fragments 5' of exon 26 were used. The extent of the 21.9-kb deletion in H51 DNA is shown below the diagram of the 3' portion of the factor VIII gene. Due to the presence of repetitive DNA near the deletion boundaries, there is an uncertainty of ~2 kb in the precise deletion end points. **b**, A Southern blot of *SstI*-digested haemophilic DNAs similarly revealed that H96 DNA lacks an 8.3-kb fragment that is present in every other sample tested. The 39.0-kb deletion in H96 was mapped using one cDNA and three cloned genomic fragments from λ 114, λ 482 and p542 (refs 2, 3).

RNA. Neither patient H2 or H22 possesses detectable levels of factor VIII activity. (It has not been possible to measure truncated protein immunologically, as patient H2 produces high levels of anti-factor VIII antibodies, and patient H22 receives frequent infusions of factor VIII for treatment.) Eventual expression of these mutant genes in a heterologous cell system² may resolve these questions.

Finally, two cases of partial deletions of the factor VIII gene were detected in our survey. A Southern blot (Fig. 4) demonstrates that the DNA of patient H51 (a severe haemophiliac but not an inhibitor patient) lacks a 3.0-kb *RsaI* fragment, which contains exon 26. Similarly, a Southern blot of *SstI*-digested haemophilic DNAs indicates that the DNA from the severe haemophilic H96, who produces antibodies (10–50 Bethesda units), lacks an 8.3-kb fragment that includes exons 23–25. The extents of the H51 and H96 deletions were mapped by hybridizing genomic Southern blots with a series of cloned genomic probe fragments at the 3' end of the factor VIII gene. In H96 DNA, the section missing from the factor VIII gene extends from 1 kb 3' of exon 22 to 17.5 kb 5' of exon 26, a total deletion of 39.0 kb. The deletion in H51 DNA, totalling 21.9 kb, extends roughly from 1 kb 5' of exon 26 to 19 kb 3' of the end of the transcribed factor VIII gene.

The 92 haemophilia A patients in this study can be categorized on the basis of their factor VIII activity as follows: 73 severe (<1% activity), 7 moderate (1–5% activity) and 12 mild (>5% activity). Fourteen of the severe haemophiliacs also produce inhibiting anti-factor VIII antibodies. All of the 92 DNA samples have been analysed with at least 2 restriction enzymes, while

all 14 inhibitor patients were extensively examined with at least 6 enzymes to ensure resolution of all 26 exons of the factor VIII gene. In contrast to the findings of Giannelli *et al.*¹⁰, who reported deletions in the factor IX gene of four out of five haemophilia B inhibitor patients, we observed only one case of a deletion (H96) in the factor VIII genes of the 14 haemophilia A inhibitor patients tested. The only other patient with a deletion found in this survey, H51, does not produce antibodies. Therefore, the suggestion that there is a correlation between the presence of antibodies and a deletion of part of the gene¹⁰ is not borne out in the case of factor VIII.

We thank Yu-Cheng Liao and Doug Smith for assistance with DNA sequencing, Karen Wion and Philip Hollingshead for help with mapping, Jeanne Arch, Alane Gray and Carol Morita for manuscript preparation; Susan Karp, Laurie Weisberg and Peter Jones for collecting blood samples from haemophilia patients, Dennis Drayna for providing non-haemophilic DNA samples, and the many patients and their families who participated in this study. This research was supported by Genentech, Inc.

Received 22 January; accepted 20 March 1985.

1. Haldane, J. B. S. *J. Genet.* **31**, 317–326 (1935).
2. Wood, W. I. *et al. Nature* **312**, 330–337 (1984).
3. Gitschier, J. *et al. Nature* **312**, 326–330 (1984).
4. Gitschier, J., Drayna, D., Tuddenham, E. G. D., White, R. L. & Lawn, R. M. *Nature* **314**, 738–740 (1985).
5. Barker, D., Shafer, M. & White, R. *Cell* **36**, 131–138 (1984).
6. Fulcher, C. A., Roberts, J. R. & Zimmerman, T. S. *Blood* **61**, 807–811 (1983).
7. Vehar, G. A. *et al. Nature* **312**, 337–342 (1984).
8. Breathnach, R. & Chambon, P. A. *Rev. Biochem.* **50**, 349–383 (1981).
9. Schmid, C. W. & Jelinek, W. R. *Science* **216**, 1065–1070 (1982).
10. Giannelli, F. *et al. Nature* **303**, 181–182 (1983).
11. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning* (Cold Spring Harbor Laboratory, New York, 1982).
12. Taylor, J. M., Illmensee, R. & Summer, S. *Biochim. biophys. Acta* **442**, 324–330 (1976).

A role for branchpoints in splicing *in vivo*

G. Rautmann & R. Breathnach

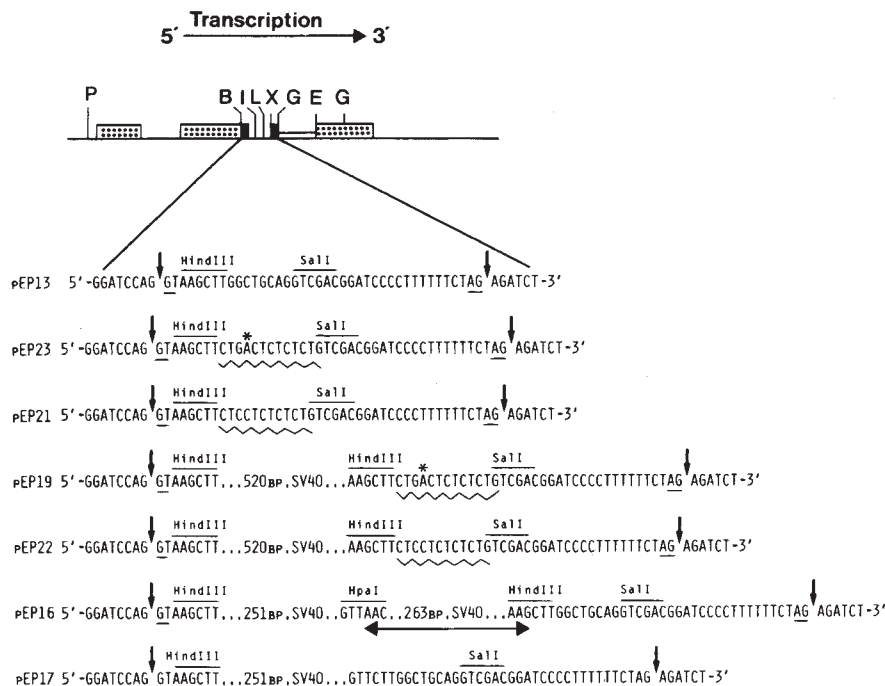
Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS, Unité 184 de Biologie Moléculaire et de Génie Génétique de l'INSERM, Institut de Chimie Biologique, Faculté de Médecine, 11, rue Humann, 67085 Strasbourg Cedex, France

The nucleotides immediately surrounding intron/exon junctions of genes transcribed by RNA polymerase B can be derived from 'consensus' sequences for donor and acceptor splice sites by only a few base changes^{1–3}. Studies *in vivo* have underlined the importance of these junction nucleotides for splicing^{4–7}. In higher eukaryotes, no evidence has been found for specific internal intron sequences involved in splicing^{8–10}. However, the recent discovery that, *in vitro*, introns are excised in a lariat form where the 5' end of the intron is joined via a 2'-5'-phosphodiester linkage to an A residue (branchpoint acceptor) close to the 3' end of the intron, suggests that internal intron sequences may nonetheless be important for splicing^{11–13}. Indeed, in yeast nuclear genes, the internal sequence 5'-TACTAAC-3' (or close homologue) is essential for splicing *in vivo*^{14,15}. A proposed consensus sequence for branchpoints in mammalian introns is 5'-CT(A/G)A(C/T)-3' (refs 11, 16, 17). This sequence resembles the essential yeast internal sequence. Are branchpoints involved in the splicing of introns of higher eukaryotes *in vivo*? We show here that a branchpoint sequence from a human globin gene (5'-CTGACTCTCTCTG-3')¹¹ greatly enhances the efficiency of splicing of a 'synthetic' intron in HeLa cells. A mutated branchpoint sequence, 5'-CTCCTCTCTCTG-3', in which the branchpoint acceptor nucleotide A has been deleted and the neighbouring purine G mutated to a C, does not exhibit this enhancing capability. We conclude that branchpoints have an important function in the splicing process *in vivo*.

We have described elsewhere the construction of synthetic splice sites designed to conform to the consensus sequences for

Fig. 1 Transcription units used to study splicing of synthetic introns. Dotted boxes represent rabbit β -globin gene exons and solid boxes represent exon sequences of the splicing oligonucleotide as described elsewhere⁷. Thus, the 5' solid box represents the sequence 5'-GGATCCAG-3' and the 3' solid box the sequence 5'-AGATCT-3'. The *Bgl*II/*Eco*RI fragment shown as a double line represents a ~100-bp fragment of pBR328 (see ref. 7 for a full discussion). P, *Pst*I; B, *Bam*HI; I, *Hind*III; L, *Sal*I; X, *Xba*I; G, *Bgl*II; E, *Eco*RI. Sequences of synthetic introns described in the text are shown. Splice points are marked by arrows, and the invariant dinucleotides GT and AG present at the 5' and 3' ends of introns, respectively, are underlined. Construction of pEP13, pEP16 and pEP17 has been described elsewhere⁷. The *Hpa*I/*Alu*I fragment of the internal intron sequence of pEP16 deleted to make pEP17 is shown by horizontal arrows below the pEP16 sequence. *Hind*III and *Sal*I sites used to introduce the globin branchpoint sequence¹¹ 5'-CTGACTCTCTCTG-3' or its mutated counterpart 5'-CTCCTCTCTCTG-3' are shown (wavy underlining). The branchpoint acceptor nucleotide A used to make a 2'-5'-phosphodiester linkage with the G residue at the 5' end of the intron is shown by an asterisk in the pEP23 and pEP19 sequences.

Methods. To introduce the branchpoint sequence, chemically synthesized¹⁸ oligonucleotides 5'-AGCTTCTGACTCTCTCTG-3' and 5'-TCGACAGAGAGAGTCAGA-3' (phosphorylated at their 5' ends by treatment with ATP and polynucleotide kinase) were mixed in a 1:1 ratio and ligated to a phosphatase-treated *Hind*III + *Sal*I digest of pEP13. pEP23 generated in this manner was linearized with *Hind*III and ligated to a *Hind*III fragment of SV40 DNA (bp 3,476–4,002, BBB numbering system¹⁹) to make pEP19. The same *Hind*III fragment is present in the same orientation in pEP16, pEP19 and pEP22. It has been represented in a different fashion in pEP16 only to show the deletion used to generate pEP17. A similar strategy was used to construct pEP21 and pEP22. Drawings are not to scale.



donor and acceptor splice sites and their expression *in vivo* as RNA from a β -globin gene vector⁷. This vector (pEP1, Fig. 2) contains a rabbit β -globin gene linked to simian virus (SV40) enhancer. In plasmids designed to express the synthetic splice sites, the *Bam*HI/*Eco*RI fragment of pEP1 carrying the globin large intron and flanking exon/intron junctions was replaced by a *Bam*HI/*Eco*RI fragment carrying the synthetic donor (5'-GGATCCAG-GTAAGCTT-3') and acceptor (5'-CTTTTCTAG-AGATCT-3') splice sites separated by various synthetic introns. In pEP13, the synthetic splice sites are separated by an internal intron sequence derived from the bacterial plasmid pUC8 (Fig. 1). In pEP16, a 526-base pair (bp) *Hind*III fragment from SV40 genomic DNA has been added to the internal intron sequences of pEP13 (Fig. 1). When introduced into HeLa cells by the calcium phosphate co-precipitation technique, both of these plasmid constructions lead to a very small amount of cytoplasmic RNA in which the synthetic donor splice site has been correctly spliced to the synthetic acceptor splice site⁷.

Deletion of some internal intron sequences from pEP16 (arrows in Fig. 1) generated pEP17 (Fig. 1). Much greater amounts of correctly spliced RNA were obtained from cells transfected with pEP17 than with either pEP13 or pEP16 (see Fig. 10 of ref. 7 and below). It seemed possible that pEP13 and pEP16 lacked a sequence capable of functioning efficiently as a branchpoint. The deletion used to generate pEP17 might have fortuitously positioned a branchpoint sequence at a suitable distance from the synthetic acceptor, thereby greatly increasing splicing efficiency. A possible candidate for such a sequence is the pentanucleotide 5'-TTAAA-3' positioned 50 nucleotides away from the acceptor splice site in pEP17 (within the 251-bp SV40 DNA segment (Fig. 1), but over 270 nucleotides away from the acceptor splice site in pEP16 (the sequences 5'-TTAAC-3' and 5'-TTAAT-3' have been proposed as branchpoints¹⁶). To test the hypothesis that pEP13 and pEP16 lack an efficient branchpoint sequence, we decided to introduce a known branchpoint sequence into these plasmids and test its effect on splicing.

The *Hind*III/*Sal*I fragments in the internal intron sequences of pEP13 and pEP16 were replaced by synthetic *Hind*III/*Sal*I fragments carrying the identified branchpoint sequence (5'-CTGACTCTCTCTG-3') of a human globin gene¹¹ to generate pEP23 and pEP19, respectively (Fig. 1; pEP16 and pEP19 contain the same *Hind*III fragment in the same orientation). Each of these four plasmids and pEP17 was introduced by the calcium phosphate co-precipitation technique into HeLa cells, together with pEP1 (to monitor the efficiency of the transfection). Cytoplasmic RNA was collected 72 h later and used for *S*₁ nuclease mapping with a mixture of two 5'-end-labelled single-stranded DNA probes. One of these probes (probe A, Fig. 2) detects correctly spliced globin RNA derived from pEP1; this RNA will protect 54 nucleotides of probe A from *S*₁ nuclease digestion. The second probe (probe D, Fig. 2) is designed to detect RNA derived from pEP13, pEP16, pEP17, pEP19 or pEP23 in which both the small globin intron and the synthetic intron have been correctly removed by splicing. Such spliced RNA will protect 325 nucleotides of probe D from *S*₁ nuclease digestion (Fig. 2).

Figure 3 shows that much less spliced RNA is obtained with pEP13 and pEP16 than with pEP17, confirming our previous observations⁷. (We term spliced RNA, here and below, as RNA in which the synthetic donor site has been spliced to the synthetic acceptor site. We have shown previously⁷ that when the synthetic acceptor is not used for splicing to the synthetic donor, splicing of the latter to a cryptic splice site in the *Eco*RI/*Bgl*II globin exon fragment occurs. This RNA cannot be detected using probe D.) Of the plasmids containing the globin branchpoint sequence, one (pEP23) yielded low levels of spliced RNA similar to those observed with pEP13 and pEP16, while the other (pEP19) yielded much higher levels of spliced RNA, similar to those observed with pEP17. It is highly unlikely that these results are due to different transfection efficiencies, as in all cases, roughly similar amounts of globin RNA were produced from the co-transfected pEP1 (as evidenced by the intensities of the ~54-nucleotide protected fragment in each lane of Fig. 3).

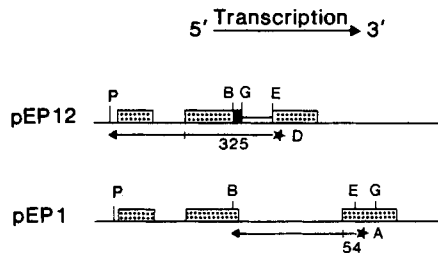


Fig. 2 Probes used for S_1 nuclease mapping. Symbols are as described in Fig. 1 legend. Construction of pEP1 and pEP12 has been described elsewhere⁷. Briefly, pEP1 contains a rabbit β -globin gene linked to an SV40 enhancer. pEP12 represents a cDNA version of pEP13 from which the synthetic intron sequence has been deleted using the enzymes *Bam*HI and *Xba*I. For preparation of probe D, a *Pst*I/*Eco*RI fragment of pEP12 was 5'-end labelled with [γ -³²P]ATP and polynucleotide kinase, and strand-separated as described previously⁷. The strand labelled at the *Eco*RI terminus was isolated. RNA extracted from HeLa cells transfected with any of the plasmids pEP13–pEP23 in which both the small globin intron and the synthetic intron have been removed correctly will protect 325 nucleotides of probe D from nuclease S_1 digestion as shown (see ref. 7 for a more detailed discussion). For preparation of probe A, a *Bam*HI/*Eco*RI fragment of pEP1 was 5'-end-labelled and strand-separated. The strand labelled at the *Eco*RI terminus was isolated. Globin RNA from pEP1 will protect 54 nucleotides of this probe from nuclease S_1 digestion as shown. Drawings are not to scale. Cuts on the arrows indicate extents of probes protected by correctly spliced RNA.

Introduction of a globin branchpoint sequence into the synthetic intron of pEP16 thus leads to a large increase in splicing efficiency. A similar effect is not seen for pEP23, perhaps because in this plasmid the branchpoint acceptor nucleotide A of the sequence 5'-CTGACTCTCTCTG-3' is too close to the 5' end of the intron (only 11 nucleotides away). However, there are clearly other differences between pEP19 and pEP23 (for instance intron size) which could explain the reduced splicing efficiency observed with the latter construction. We also show in Fig. 3 results obtained with pEP18, a version of pEP23 containing a branchpoint of the form 5'-CTGATCTCTCTG-3' and thus carrying a single-nucleotide deletion relative to the pEP23 branchpoint. A low level of spliced RNA, similar to that seen with pEP23, was obtained with pEP18. If the significantly higher efficiency of splicing observed with pEP19 as compared with pEP16 is due to the presence of an efficient branchpoint sequence in the former, elimination of the branchpoint acceptor nucleotide A from the sequence 5'-CTGACTCTCTCTG-3' present in pEP19 should greatly decrease splicing efficiency. The plasmid pEP22 contains the mutated branchpoint sequence 5'-CTCCTCTCTCTG-3' (Fig. 1). For the mutated branchpoint sequence, we chose to delete the branchpoint acceptor nucleotide A and to mutate the neighbouring G residue to a C, as, although the branchpoint acceptor nucleotides identified to date are all A, it has not been demonstrated that a G cannot function as a branchpoint acceptor. We hypothesized that a pyrimidine would be less likely to function in this way, and therefore mutated the G to a C. Levels of spliced RNA obtained with pEP22 were indeed low (Fig. 3), and comparable to those obtained with pEP13. As expected, pEP21, a version of pEP23 with the mutated branchpoint sequence, also gave rise to the low level of spliced RNA apparently characteristic of constructions lacking an efficient branchpoint sequence.

We have shown that introduction of a globin branchpoint sequence into a synthetic intron can greatly enhance splicing efficiency, thus demonstrating the importance of branchpoint sequences for splicing *in vivo*. We assume that the low levels of spliced RNA obtained with constructions such as pEP13 and pEP16 reflect the use of an inefficient branchpoint, although it has not been shown that a branchpoint is essential for splicing.

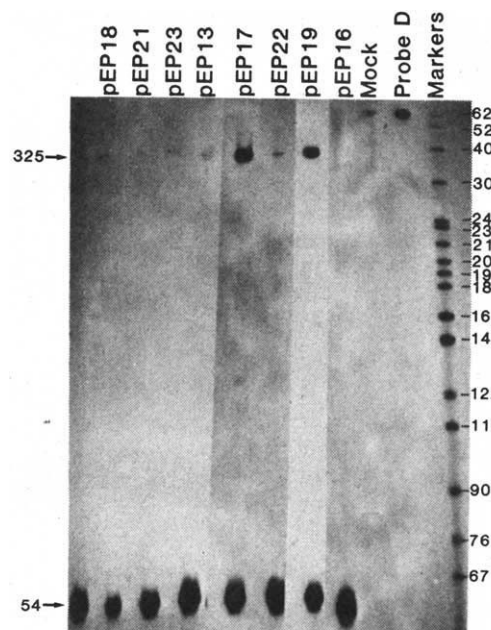


Fig. 3 Effect of a globin branchpoint sequence on splicing of synthetic introns. Plasmids shown in Fig. 1 were mixed in a 4:1 ratio with pEP1 and introduced into HeLa cells by the calcium phosphate co-precipitation procedure. Cytoplasmic RNA was isolated 72 h later and 50 μ g used for nuclease S_1 mapping with a mixture of probes A and D (Fig. 2). Protected fragments of probe were electrophoresed on an 8% polyacrylamide 7 M urea gel. Size markers are from an *Msp*II digest of pBR322. Fragments of probe A protected by RNA from pEP1 are indicated by their size, ~54 nucleotides. Fragments of probe D protected by RNA from pEP13–pEP23 are also indicated by their size, ~325 nucleotides. Mock, mock-transfected HeLa cells. All experimental conditions were as described previously⁷.

The sequence features defining a branchpoint and its optimal position relative to an acceptor splice site are unknown. Plasmids such as pEP16 which lack an efficient branchpoint sequence should prove useful in this respect.

We thank P. Chambon for useful discussions, M. C. Gesnel, M. Acker and B. Augsburger for technical assistance, A. Staub for synthesis of the oligonucleotides used in this work, and B. Boulay, C. Werlé and C. Aron for preparation of the manuscript. G.R. was a recipient of a grant from the Ligue Nationale Française contre le Cancer. This work was supported by grants from the CNRS, INSERM, the MIR (88V0092), the Fondation pour la Recherche Médicale and the Association pour le Développement de la Recherche sur le Cancer.

Received 16 January; accepted 21 March 1985.

1. Breathnach, R., Benoist, C., O'Hare, K., Gannon, F. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4852–4857 (1978).
2. Breathnach, R. & Chambon, P. *A. Rev. Biochem.* **50**, 349–383 (1981).
3. Mount, S. *Nucleic Acids Res.* **10**, 465–473 (1982).
4. Wieringa, B., Meyer, F., Reiser, J. & Weissmann, C. *Nature* **301**, 38–43 (1983).
5. Treisman, R., Orkin, S. H. & Maniatis, T. *Nature* **302**, 591–596 (1983).
6. Mount, S. & Steitz, J. *Nature* **303**, 380–381 (1983).
7. Rautmann, G., Matthes, H. W. D., Gait, M. J. & Breathnach, R. *EMBO J.* **3**, 2021–2028 (1984).
8. Volkaert, G., Feunteun, J., Crawford, L. V., Berg, P. & Fiers, W. J. *J. Virol.* **30**, 677–682 (1979).
9. Wieringa, B., Hofer, E. & Weissmann, C. *Cell* **37**, 915–925 (1984).
10. Dierks, P. et al. *ICN UCLA Symp. molec. cell. Biol.* **33**, 347–366 (1981).
11. Ruskin, B., Krainer, A. R., Maniatis, T. & Green, M. R. *Cell* **38**, 317–331 (1984).
12. Padgett, R. A., Konarski, M. M., Grabowski, P. J., Hardy, S. F. & Sharp, P. A. *Science* **225**, 898–903 (1984).
13. Weissmann, C. *Nature* **311**, 103–104 (1984).
14. Pikielny, C. W., Teem, J. L. & Rosbash, M. *Cell* **34**, 395–403 (1984).
15. Langford, C. J. & Gallwitz, D. *Cell* **33**, 519–527 (1983).
16. Keller, E. B. & Noon, W. A. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7417–7420 (1984).
17. Zeitlin, S. & Efstratiadis, A. *Cell* **39**, 589–602 (1984).
18. Matthes, H. W. D. et al. *EMBO J.* **3**, 801–805 (1984).
19. Toozé, J. (ed.) *DNA Tumor Viruses* 2nd edn (Cold Spring Harbor Laboratory, New York, 1982).

Most curious is his endorsement of the view (first stated by Derek Price in 1957) that Copernicus was no revolutionary. It seems odd to minimize the cosmological and epistemological significance of Copernicus; if the Pythagoreans had long preceded him, Newton (with less occasion, we may think) allowed as much with respect to the theory of gravitation, for he thought that the Pythagoreans must have understood this also. No greater tribute could be paid to Copernicus than that implied in Newton's conclusion to the first Scholium of *Principia Mathematica* (1687): I wrote the following treatise, he declared, with the object of elucidating true motions and causes in nature from apparent causes and effects, and conversely of elucidating causes and effects from both true and apparent motions. If Copernicus had not proposed this problem, who did?

The Newtonian system without Copernicus as the first of the "giants" is indeed inconceivable. Bernard Cohen's difficulty arises from the perhaps inavoidable confusion between the idea of revolution as the work of an individual (Lenin, Lavoisier) and the idea of revolution as the fruition of a movement (bolshevism, pneumatic chemistry). Sometimes Cohen seems virtually to be defending the "great man" concept of change, or even (in his discussion of Wegener) the "neglected genius". More frequently, he seems to mark the great men on a scale of revolutionary activity or fervour: thus Bacon's great contribution to methodology "did not constitute a revolution in science"; Gilbert "did not establish a new science", unlike Galileo, "a heroic figure of the Scientific Revolution", although even "Galileo's revolution was not complete". Descartes's revolution "did not last" and "it seems clear there was no Keplerian revolution in science before 1687". One remembers that Alfred was a Good King. What is historically important is an understanding of how the separate, even disparate, endeavours of individuals contemporaneous with each other or in quick succession, culminate in some particularly noteworthy and well-founded shift in our vision of nature, for few historians believe that scientific revolutions are literally effected single-handed by a unique scientist. And so to me, for example, it makes good sense to speak of a "revolution in geology" between about 1790 and 1840, whereas talk of a "Lyellian revolution" would be unjustifiable. Or again, it is more constructive to consider the triumvirate Faraday-Maxwell-Hertz as a movement, than to try to weigh out the praise for innovation due to each.

Like many large historical works, Bernard Cohen's book excels as a quarry, rather than as a coherent argument. He is likely to have the last word for a long time to come. □

A. Rupert Hall is Emeritus Professor of the History of Science and Technology at the University of London.

Agents of change

S. Venitt

Chemical Carcinogens, 2nd Edn. ACS Monograph 182.

Edited by Charles E. Searle.

American Chemical Society: 1984. Pp. 1,373. \$129.95 (North America), \$155.95 (elsewhere).

UNTIL recently, the study of carcinogenesis has been largely correlative and phenomenological. What does chemical A do to biological systems X, Y and Z? Does B do it better and is it metabolized to C? What happens when you apply A before B, or B before A, or mix A and B together? And so forth. This approach has been very fruitful: we now know, for example, that most carcinogens are genotoxic and that most non-carcinogens are not; that many carcinogens are themselves electrophiles or are metabolized to electrophiles; that some substances "promote" tumours but do not initiate them; that many carcinogens are organotrophic; and that carcinogenesis is a multi-step process. Much of this information was gathered together in the first edition of *Chemical Carcinogens*, published in 1976.

The study of the mechanism of carcinogenesis was transformed and became at once both analytical and synthetic when two groups independently showed that NIH-3T3 cells could be transformed *in vitro* by transfection with DNA from a human bladder tumour, and that a single base change within a proto-oncogene was in some cases sufficient to activate the oncogene. There has followed a flood of research (which shows no sign of abating) which stresses the importance of both point mutation and chromosomal rearrangement in the activation of proto-oncogenes. The idea that DNA is the critical target for carcinogens has thus gained support from two different streams: the older, painstaking, sometimes stumbling correlative approach, and the rather more breezy and with-it analytical methods of cell biology and recombinant DNA technology.

Nevertheless, "descriptive chemical carcinogenesis" will not go away. We still need to know whether chemicals, new or already in use, are likely to be carcinogenic. In this light appears the second edition of the American Chemical Society's *Chemical Carcinogens*, which offers an expanded view of the descriptive approach embodied in the first edition.

ACS *Chemical Monographs* are intended to serve two main functions: (i) "to make available to chemists a thorough treatment of a selected area in form usable by persons working in more or less unrelated fields to the end that they may correlate their own work with a larger area of physical science"; and (ii) "to stimulate further research in the specific field treated". The first edition filled these roles

admirably, not only for chemists, but also for a wider audience interested in chemical carcinogenesis. This second edition, in two well-indexed and richly-referenced volumes, is a valuable update and expansion: there are now 22 chapters (all by leading specialists in each field) instead of 16.

Included are completely new chapters on cancer epidemiology, DNA-carcinogen-metabolite interactions, halogenated organic compounds, inorganic carcinogens, environmental *N*-nitroso carcinogens, triazenes, hydrazines, azo and azoxy compounds, aflatoxins, fusarial mycotoxins, carcinogenic medicines, and the inhibition of chemical carcinogenesis. Other topics which were covered in separate contributions in the first edition (for example, metabolism of carcinogens, carcinogens in plants and micro-organisms) are now spread over several. Although much of the core of the first edition remains (polynuclear aromatic hydrocarbons; soots, tars and oils; aromatic amines and their epidemiology; laboratory hazards; alkylating agents; mineral fibres; *N*-nitroso compounds; bracken carcinogens; carcinogens in food; bioassays and short-term tests) several important topics are omitted, such as tumour promotion and co-carcinogenesis, endocrine aspects of carcinogenesis and respiratory carcinogenesis. This is a pity, and it is fair to assume that the vagaries of deadlines and non-availability of authors dictated the final content.

Much of the material consists of detailed and very helpful tabulations of carcinogens and descriptions of their metabolism and biological activity, with much emphasis on structure-activity relationships. In a text of this size and scope there are bound to be some contributions which are more thorough or more readable than others and it is tempting to give each of them marks out of ten. Although I have resisted the urge I must nevertheless single out Lawley's contribution on carcinogenesis by alkylating agents as the chapter which gives the deepest and broadest view of chemical carcinogenesis as it stood in 1981-1982. Viewing alkylating agents as "archetypal carcinogens", he provides us with a masterly, wide-ranging review, which includes discussions of chemical reactivity, dose-response relationships, mechanisms of mutagenicity, short-term tests, DNA repair, models of carcinogenesis, alkylating agents and human cancer — and, as a final flourish, squeezes in (in a note added in proof) a short account of the first reports of activation of proto-oncogenes by point mutation.

Orchestrating such an encyclopaedic text must have been an awesome task. The editor, Dr C.E. Searle, must be congratulated for producing such a useful revision of an already indispensable work. □

S. Venitt is a member of the Section of Chemical Carcinogenesis, Institute of Cancer Research, Royal Cancer Hospital, Sutton, Surrey.

Faulting the heavens

David W. Hughes

Under Newton's Shadow: Astronomical Practices in the Seventeenth Century.

By Lesley Murdin.

Adam Hilger: 1985. Pp. 152. £16.75, \$26.

WHAT was it like being an astronomer in the Newtonian age (1643–1727)? There is no easy answer to the question as there was in fact only *one* full-time professional and even he, the Reverend John Flamsteed, diverted himself with his parish at Burstow, instrument-making and tide tables. The remainder were scientific generalists who saw little reason to specialize and flitted from subject to subject apparently as the whim took them. Even Edmond Halley, probably next in line for the astronomer appellation, spent only a quarter of his time studying lunar motion, planetary transits, occultations, eclipses, southern stars, proper motion and comets. For the rest he enjoyed the pleasures of meteorology, geomagnetism, physics, pure mathematics, and the editing and translating of ancient mathematical texts, not to mention his duties at various times as a captain in the Royal Navy, a diplomat and the deputy comptroller of the Chester Mint.

The Newtonian age was a time of transition. The pre-Civil-War concept of a magical and mystical world was tackled by men with a desire to avoid speculation and a preference for observation and the accumulation of data. Aristotelian views of a changeless, perfect (and extremely boring) heaven crumbled with the invention and development of telescopic instruments. Astronomers were provided with a goal, the discovery of an easy method of finding the longitude of a ship at sea, and the government were even prepared to spend money to find an answer. In the new Royal Society the scientific generalists found an excellent means of exchanging ideas, meeting like-minded friends and publishing their results. The gap between amateur and professional astronomer was only just beginning to open up. All that was needed was a grammar school education, enthusiasm, a cheap refracting telescope, an altitude measurer and a clock.

Lesley Murdin has tried to introduce us to the sort of men who found astronomy fascinating. The list is mercifully short — Flamsteed, a dozen or so university dons (for example Gregory, Halley, Hooke, Newton, Wallis, Ward, Whiston and Wren) who spent a fraction of their time with stars and planets, another dozen interested amateurs (among them Towneley, Derham, Pound, Gray and Moore) who had also to work for a living, and a final dozen “observatory technicians” (Crosthwait, Denton, Leigh and Sharp, for instance). From this narrow base Murdin has drawn a series of sweeping conclusions. “The number of astronomers who never

married seems high and might suggest a general tendency to put work before relationships.” Astronomers were independent thinkers and creative but generally had “an interest in objects rather than people” and “found difficulty in co-operating with others”. They exhibited “a tendency to solitariness”.

At the end of the book I was left with a very sketchy idea of the motivation of astronomers as such, and of their specific problems of learning, observing, calculating, interrelating, publishing, teaching and instrument-making. Little is

said about their standing in the community. Murdin has provided us with a readable gallop through the astronomical seventeenth century but one which makes clear that much more work needs to be done in this field. Additionally, the lack of references is frustrating, often leaving the reader with no idea of the provenance of a quotation. An index that only lists proper names also hints at a job half done. □

David W. Hughes is Senior Lecturer in Astronomy and Physics at the University of Sheffield.

Diverse questions

Elvin A. Kabat

The Antibody Enigma.

By Thomas J. Kindt and J. Donald Capra.

Plenum: 1984. Pp. 270. \$35, £33.25.

A PERSON wishing to learn about the structural and genetic basis of antibody diversity and antibody specificity might well be overwhelmed by the literature. *The Antibody Enigma* is a good place to start to get to grips with the subject.

The authors do not follow a strictly historical treatment but attempt to present the subject in the light of the working hypotheses considered and discarded as new experimental data became available. As they state in the preface, “*The Antibody Enigma* is a somewhat personal view of the antibody diversity problem from two investigators who have spent the past 18 months trying to penetrate the enigma”. Presumably, much of this was done in face-to-face discussion, to formulate points of agreement and resolve differences. I wish I could have listened in. It would have been difficult not to interrupt, however, for despite enormous progress this is an area which is still filled with differences of opinion, of interpretation, and of judgement about the relative weight to be given to the various mechanisms contributing to the vast repertoire of antibody combining sites.

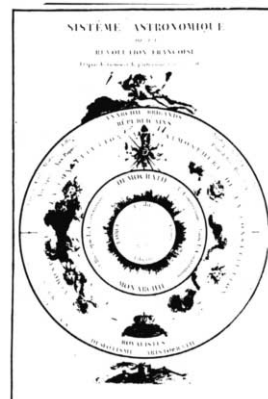
After defining the problem, the authors set down the serologists’ and biochemists’ approaches to it and their unique features. In subsequent chapters, they analyse “polar”, specifically the germ-line versus somatic mutation hypotheses, and the “maverick” solutions (those based on data not in accord with either). This is followed by discussion of the contributions of molecular biology and a final chapter entitled “Antibody Diversity: A Contemporary Solution”. Many key references are provided, but the authors often mention individuals by name with no reference and at many points I found it frustrating not to be able to pin down the source of some of the statements.

In the preface, Kindt and Capra apologize if they have “offended anyone by

omitting their own contribution...”; given the vastness of the literature they may well have. In line with the general trend among molecular biologists who have established so many ways of generating antibody diversity that they no longer consider how much of this diversity is related to antibody complementarity, the authors do not deal with this question. More understandably, there is no hint of the ferment of activity consequent upon the cloning and sequencing of the T-cell receptor (the book was in press when all the excitement began). But clearly this is a volume that should be read by individuals wishing to enter the field. □

Elvin A. Kabat is Higgins Professor of Microbiology and Professor of Human Genetics and Development at Columbia University.

—Advertisement—



Courtesy of Henry E. Huntington Library & Art Gallery

Revolution in Science

I. BERNARD COHEN

Spanning five centuries and virtually all of scientific endeavour, *Revolution in Science* seeks to uncover nothing less than the nature of *all* scientific revolutions, the stages by which they occur, their time scale, specific criteria for determining whether or not there has been a revolution, and the creative factors in producing a revolutionary new idea. Weaving threads of detail from physics, mathematics, Freud, atomic physics and even plate tectonics and molecular biology, the book is a probing analysis of the history of an idea and one of the most impressive surveys of the history of science ever undertaken.

Belknap £22.75 Hardback 704pp with 1 colour, 6 b/w halftones & 6 line drawings 0-674-76777-2

HARVARD

UNIVERSITY PRESS

126 Buckingham Palace Road London SW1W 9SD

Tissue culture equipment

New items available for the tissue culture laboratory include carbon dioxide incubators, novel bioreactors and media.

● Imperial Laboratories of Salisbury, working with the Wessex Regional Cytogenetics Unit at Salisbury General Hospital, have developed a new medium for use in cytogenetics procedures. This medium, based on Iscove's modification of Dulbecco's MEM, has the folic acid level reduced from the normal 4.00 mg to 0.01 mg per litre. The folate-deficient Iscove's medium has proved successful in producing mitotic indices, with well-spread elongated chromosomes suitable for high resolution banded analysis, and is now used routinely at the Salisbury hospital for all lymphocyte cultures and also for bone marrow specimens from patients with leukaemia. *Reader Service No. 100.*

● A new cell culture medium, Ventrex PC-1, is now available from Tissue Culture Services to complement the Ventrex HL-1 medium. Like HL-1, Ventrex PC-1 is a low-protein, serum-free medium which has been specially formulated using defined components for optimal cell growth, while maintaining the lowest possible protein content (less than 750 µg per ml). The PC-1 medium is intended for the culture of primary cells and cell lines and has been successfully tested on many widely used cell lines. Ventrex PC-1 is available as a sterile liquid base medium, with a sterile frozen supplement, in 1 × 100 ml and 2 × 500 ml packs. When stored under suitable conditions, the components are stable for 120 days; when the components are combined to form the complete medium, it remains stable for 45 days at 4°C. The medium is also available in a 10-litre component package for users who prefer to prepare their own media. *Reader Service No. 101.*

● Opti-MEM I from Gibco is a multi-use reduced serum medium which supports optimal growth with reduced fetal bovine serum for a wide variety of mammalian cells. Most cells routinely cultured in serum-supplemented medium may be directly transferred into Opti-MEM I with a minimum of 50% reduction in serum. Opti-MEM I is performance-tested for cloning efficiency and growth promotion using various mammalian cell types, including murine myelomas, hybridomas and adherent cultures. Opti-MEM I has a shelf life of six months from date of manufacture, and is easily prepared in only three steps: add 2-mercaptoethanol (included); mix thoroughly; and add fetal bovine serum.

These notes are based on information provided by the manufacturers. Use the reader service card bound inside the journal to obtain further information.



Clandon's new cell counter.

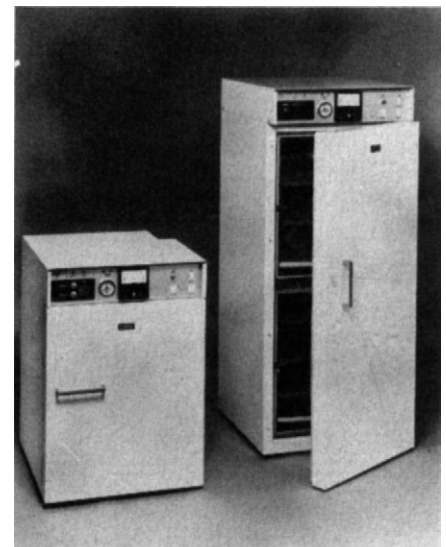
● The range of Diffcount II microprocessor-controlled differential cell counters has been extended with the introduction of Comp-U-Diff, a direct differential data entry terminal. Manufactured by Modulus Data Systems, and available in the UK from Clandon Scientific, the Diffcount II/Comp-U-Diff allows efficient, direct, computer entry of the differential counts, RBC morphology, MISO ID code, specimen identification and general comments. Straightforward instructions and suggested software enable Comp-U-Diff to be readily interfaced with an existing laboratory computer. *Reader Service No. 103.*

● MBR Bioreactor AG has introduced a membrane bioreactor for the cultivation of mammalian, plant, hybrid cells and micro-organisms. The Membroferm consists of a medium chamber, a cell chamber and a product chamber which are fitted with a special net support which acts as a carrier to fix the cells as well as supporting the membranes. This membrane stack is held in position by stainless steel endplates which are fitted with connections to the conditioning vessel where the medium can be continuously treated to maintain the correct environment for growth of the cells. Cells are introduced into the cell chamber which is separated from the medium chamber by a membrane on one side and the product chamber with a different membrane on the other side. This allows the cells to receive nutrients and gases via the medium chamber and to remove the products of metabolism or toxins via the product chamber. The technique has achieved very high cell densities with mammalian cells. *Reader Service No. 104.*

Reader Service No. 104.

©1985 Nature Publishing Group

● Plant diseases can cause crop losses ranging from slight yield depression to complete crop failure. To counter this Twyford Plant Laboratories offers 'disease indexing' services to plant breeders, propagators, seed producers and growers. Using this service it is possible to 'clean-up' a plant or crop, for further use as a mother stock for further propagation, or as parent plants for high quality seed production, for instance. TPL's indexing service is based on the use of plant tissue culture. Viruses are identified by a range of techniques including ELISA, PAGE and occasionally electron microscopy. Fungal diseases are detected using selective nutrient media. For more details on TPL's range of techniques based on an expertise in plant tissue culture use the reader service card. *Reader Service No. 105.*



Tissue culture incubators from LEEC.

● In the tissue culture laboratory, CO₂ incubators with high humidity levels are constantly subjected to contamination by yeasts, moulds and all forms of airborne bacteria. The new LEEC range of automatic CO₂ incubators incorporates a self-sterilizing facility which by-passes the normal heater controls and holds the temperature at 90°C. On completion of a cycle the control is transferred back to the normal controllers and the incubator rapidly re-stabilizes at pre-set temperatures. The range includes models of 150 litres and 320 litres capacity, each featuring dual proportional voltage controllers, safety overrides, heated inner glass door, solid state CO₂ detection and a self-sterilizing facility. The LEEC range of incubators is manufactured with high-grade stainless steel throughout. *Reader Service No. 106.*

Reader Service No. 106.



The Dynatech Autodiluter with features aimed at tissue culture.

● Through the use of modular, solid-state logic controls, the Dynatech Autodiluter III can be programmed to carry out such diverse procedures as ELISA washing and aspirating, haemagglutination, complement fixation, microtitration and a variety of tissue culture applications. The Autodiluter III can be linked to Dynatech's Dynadrop multi-reagent dispenser, for simultaneous addition of up to 12 different reagents to a row of microtitre plate wells. Aspirator dwell time can be controlled between 2 and 6 s and the downward travel of the head limited, a feature necessary in tissue culture where the Microdiluter and aspirator tips must not touch the bottom of wells.

Reader Service No. 107.

Advertisement

Agar...automatically from powder to plates in less than 90 minutes...



**Up to 8 liters
(400 plates) of
high-quality
agar whenever
you need it**

Any laboratory can have high quality agar whenever it's needed thanks to two benchtop instruments — the Agar-Matic, an 8 liter sterilizer, and the PourMatic, a self-stacking carousel-type petri dish filler.

Simply add water and dehydrated medium to the Agar-Matic. Then set it and forget it. This automatic sterilizer does the rest. Fresh, chemically superior agar or nutrient broth is ready to pour in less than an hour. If you're not back in time, the agar will be maintained at the desired pouring temperature.

By adding the PourMatic, agar can be sterilized, cooled, poured into plates and stacked... all automatically in only 38" of bench space while you tend to more important activities. Microprocessor control assures dependability and minimizes operator supervision.

Agar sterilizers from 3 to 130 liters are also available. Send for 16-page catalog. New Brunswick Scientific Co. Inc., P.O. Box 986, Edison, N.J. 08818, (201) 287-1200 or (800) 631-5417.



NEW BRUNSWICK SCIENTIFIC

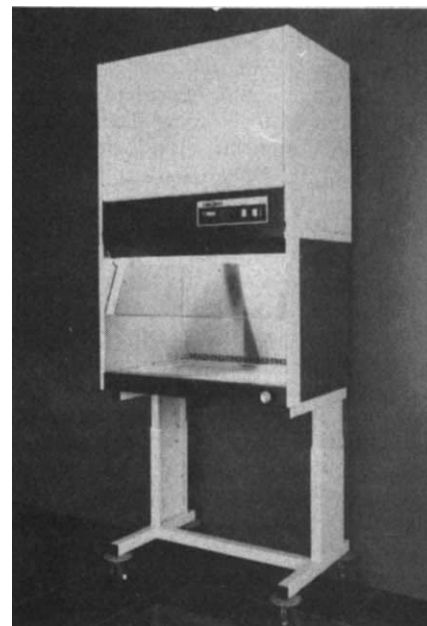
At NBS Advanced Technology is a Way of Life

Reader Service No. 49

©1985 Nature Publishing Group

● Dystonia is a neurologic disease consisting of sustained twisting movements and postures. It may be inherited, sporadic, or secondary to other nervous system diseases. The Dystonia Clinical Research Center at Columbia University has established a Fibroblast Tissue Resource Facility. Fibroblasts from patients with genetic and sporadic dystonia (as well as controls) are cultured and stored in liquid nitrogen; samples are available to investigators with established probes for biochemical markers. The Columbia centre invites scientists to utilize the tissue bank without charge. If interested, investigators should send a brief description of assays to be carried out to: Stanley Fahn, Director, Dystonia Clinical Research Center, Neurological Institute, 710 West 168th Street, New York, New York 10032, USA.

Reader Service No. 108.



Labconco's new laminar flow cabinet.

● Labconco's new range of laminar flow cabinets — the Purifier Class II — is well suited to tissue culture manipulations including inoculation of viruses. A stainless steel working area and coved and dished work surface contribute to operator protection. Other features include a variable speed motor, easy-to-read differential pressure gauge monitor, Hepa filters readily accessible for replacement, adjustable height and position sash for comfort, remote controlled gas fixture with recessed nozzle and safety valve, and adjustable height stand on heavy duty rubber-tyred castors. All lights are recessed and accessible from the cabinet exterior. The air movement and positioning of the filters places virtually all the cabinet under negative pressure, those areas of positive pressure are surrounded by areas of negative pressure ensuring complete containment in the event of a leakage. The Purifier meets all standards for Class II operation and is supplied with full performance certificates.

Reader Service No. 109.



The Bellco μ -carrier spinner flask.

● Bellco Biotechnology has introduced a new line of non-heating, precision magnetic stirrers. The Sci/ERA range is offered in one-, two- and four-position units, each featuring a port to accept computer-generated commands for external control. Stirring speeds are set with a precision ten-turn control, and are ramped up and down to minimize disruption of cell attachment. Two LED displays are incorporated; one indicates stirrer r.p.m., and the other reads out cycle and countdown to show exactly where the system is in relation to internally-controlled cycles. Electronic on/off cycle timers permit precise repetitive cycling. Stepper motors are controlled by pulses generated with such precision that repeatability of settings is ± 0.1 r.p.m., 12-volt signal level assures operation at selected speed if the system switches to back-up battery power. In addition, the new Bellco system offers a new stand-by Sci/ERA Power Pack to protect stirrers and valuable cell cultures in the event of failure of the main power supply.

Reader Service No. 110.

● The new 50-watt Model MS-50 Microson miniature ultrasonic cell disruptor is intended primarily for cell disruption, homogenization, and emulsification in the laboratory and classroom. The Microson, from Heat Systems-Ultrasonics, Inc., can also degas liquids, accelerate chemical and physical reactions, and deagglomerate and disaggregate. The unit features a new on/off momentary control switch, and micro cup horn for disruption in total isolation and a micro-Pulsar cycle timer for limiting temperature build up in heat-labile materials.

Reader Service No. 111.

● An eight-page brochure from Beckman describes two compact Microfuge centrifuges used for spinning down small samples fast. The Microfuge 11 and 12 provide a choice of three rotors. The Microfuge 11 has a fixed horizontal rotor; and the Microfuge 12 a bowl rotor. Both can spin the 13.2 fixed angle rotor and the capillary tube rotor.

Reader Service No. 112.

● A modification kit is now available to convert the standard Marubishi fermenter for the growth of cells in continuous culture. The modification is suited to both bacterial and cell cultivation units. All components are interchangeable and do not require additional expense or modifications when different size vessels are used. The standard bench-top fermentor is equipped with temperature, agitation and aeration controls. The transfer of medium is accomplished by pumps which automatically remove medium to be processed and add back medium from a reservoir.

Reader Service No. 113.

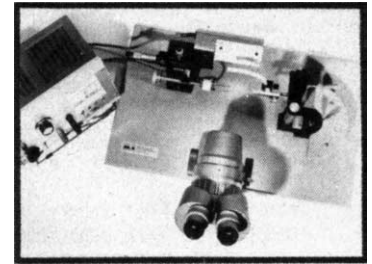


ACAS 470 workstation.

● Meridian Instruments' ACAS 470 workstation combines the methods of automated clonal isolation and fluorescence analysis into a single system that will (1) automatically isolate a clonal population of any cell type in monolayer culture; (2) perform selection and analyses of cells without removal from growth surfaces or altering morphology, phenotype or variability; (3) maintain the selected cell(s) under controlled growth conditions; (4) catalogue cells by location on a growth surface; (5) perform the full range of steady-state fluorescence analyses. Cell selection and separation are achieved by growing cells on specially treated cover slips in a standard tissue culture dish. The cover slips can be transferred to the environmental cell chamber to maintain selected cell(s) in a state of clonal growth. This autoclavable mini-incubator mounts on an inverted microscope. The cells are maintained at the desired temperature while sufficient nutrients are provided by a controllable continuous flow of medium across the inserted growth plate. A removable top allows convenient access to the cells for *in situ* manipulations.

Reader Service No. 114.

SYSTEMS FOR MICROPIPETTE PREPARATION AND MICROMANIPULATION



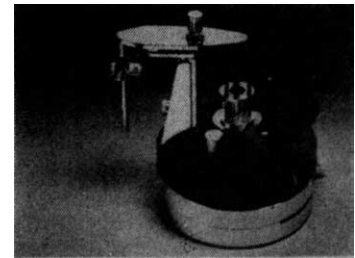
- * Microforgers
- * Micrometer syringe
- * Micromanipulators

Ideal for intracellular and intranuclear injection, embryo transfer, isolation of single cells, etc.

The premier range of high quality instrumentation for micromanipulation.

RESEARCH INSTRUMENTS LTD
Penryn, Cornwall TR10 9DQ, England
(0326) 72753.
Ebtac Corp, Mass 01001, USA
(413) 786 9322.

SINGER MICROMANIPULATOR MK III Patented



Perfect control in 3 Dimensions from a single joystick. Continuous variable reductions to suit all magnifications. No after sales maintenance needed.

Details from:
Singer Instrument Co. Ltd.,
Treborough Lodge Roadwater,
Watchet Somerset TA23 0QL
England. Tel. (0984) 40226 Telex:
337492 Comcab G.

LIPOSOMES NEW! reproductive — homogeneous — extremely rapid —

The new commercial LIPOSOMAT provides reproducible preparation of uniformly sized unilamellar liposomes (formation within 10 minutes) with very high encapsulation efficiency for both hydrophilic and/or lipophilic agents. Liposome size selectable between 25 and 60nm in diameter. Lipid loss far below 1%.

Sample volume normally 5-10 ml or up to 200 ml if used in combination with an additional apparatus.

The LIPOSOMAT® is used especially for studying drug/protein/cell — liposome interactions *in vitro* and *in vivo*, for reconstitution experiments, and for the transfer of genetic material in gene and biotechnology.

**DIANORM-Geräte, POB 126, D-8000
MUNICH 65, FRG.**

Advertisements

AUTHORS WANTED BY N.Y. PUBLISHER

A well-known New York subsidy book publisher is searching for manuscripts worthy of publication. Fiction, non-fiction, poetry, juveniles, travel, scientific, specialized and even controversial subjects will be considered. If you have a book-length manuscript ready for publication (or are still working on it), and would like more information and a free booklet, please write:

Vantage Press, Dept. D-69
516 W. 34th St., New York, N.Y. 10001

Reader Service No.44

**PHORBOL
ESTERS**

Cat. #M4600 5mg/\$55

MEZEREIN

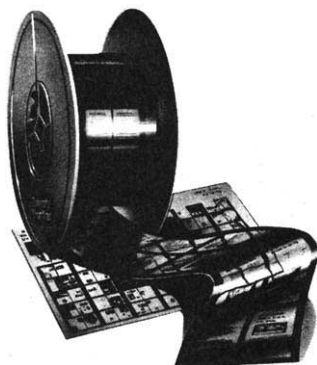
Co-Inducer of
Interferon

LC SERVICES CORPORATION

165 New Boston Street
Woburn, MA 01801
(617) 938-1700

Reader Service No.45

nature is
available in microform
from University
Microfilms
International.



☐ Please send information about these titles:

Name _____

Company/Institution _____

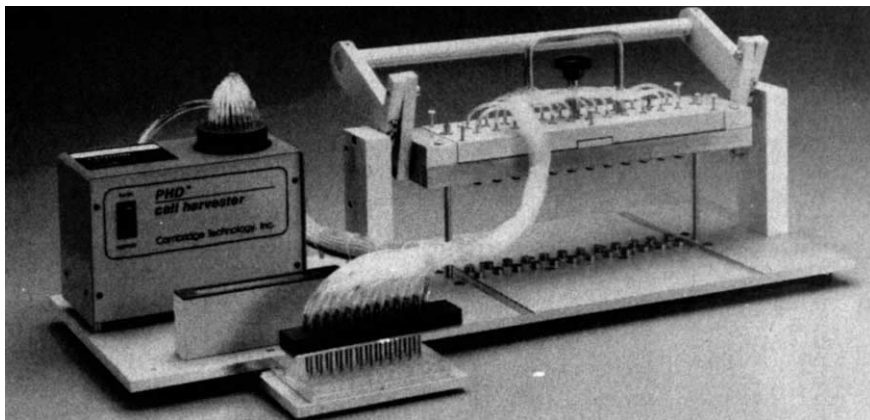
Address _____

City _____

State _____ Zip _____

Phone () _____

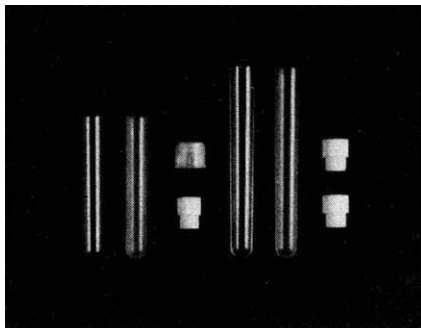
Call toll-free 800-521-3044. In Michigan,
Alaska and Hawaii call collect 313-761-4700. Or
mail inquiry to: University Microfilms International,
300 North Zeeb Road, Ann Arbor, MI 48106.



The latest version of the Cambridge Technology cell harvester.

● A new version of Cambridge Technology's PHD Cell Harvester, the Model 290, gives the user the choice of punching and depositing the samples into scintillation vials at the time of harvesting or at some later time. This feature is particularly useful when processing a large volume of samples. It minimizes the number of scintillation vial trays needed and permits compact temporary storage of harvested samples on filter mats. Either mode of operation is selected simply by turning a knob. In the non-punching mode, small keying holes are made in the filter mat during harvesting that precisely align the mats for automatic punching and depositing later. The holes also prevent the samples from being mixed-up or reversed. The unit harvests and/or punches 24 samples in each operation. All PHD accessories and options such as special heads, supernatant collection systems, wash manifolds, and quick disconnects can be used with this new instrument.

Reader Service No. 115.



Culture tubes from Bio-Rad.

● Made without mould-release agents, Bio-Rad's 13 × 100 mm culture tubes are devoid of chemical residue to interfere with analytical results. The tubes are precision-moulded to close tolerances under conditions of extreme cleanliness. Scratched or blemished tubes are eliminated by 100% inspection. The culture tubes are not sterile, but are made of polypropylene and so may be safely autoclaved. The 13-mm stoppers fit tightly and prevent leaks. All items are packaged to eliminate dust and are kept in stock for immediate shipment.

Reader Service No. 116.

● An automatic CO₂ incubator recently introduced by VWR Scientific includes a novel new control system. The Series 1800 IR automatic CO₂ incubator is equipped with an infrared control system capable of sensing a change in concentration as small as 0.1% CO₂. Available in single-chamber and dual-chamber styles, all Series 1800 IR incubators have a readable, easy-to-set CO₂ controller. A CO₂ alarm alerts the user to high or low CO₂ concentration, extra-large water jackets ensure temperature uniformity, and an access port allows easy verification of CO₂ levels.

Reader Service No. 117.

● A new serum-free medium from Costar Biologicals has been developed primarily for culturing hybridomas and T cells. This medium, SF-1, has a total protein concentration of less than 50 µg ml⁻¹ and obviates the need for addition of serum completely. Antigen-specific cytotoxic and helper T-cells can be isolated and propagated in SF-1 without loss of function. The purification of antibody-containing supernatants is greatly facilitated. SF-1 is packaged in 'normal' 100 and 500ml bottles and is stable for 6 months at +4 °C. Also new from Costar Biologicals are two hybridoma growth factors capable of eliminating the requirement for feeder layers. Both ESG (Ewing sarcoma growth factor) and HECS (human endothelial culture supernatant) enhance the efficiency of fusion and subsequent cloning (up to 80%) and maximize poliferation rate.

Reader Service No. 118.

● The typing or classification of a monoclonal antibody's immunoglobulin class or subclass is best done when the antibody is in culture supernatant. Conventional methods of typing (such as traditional Ouchterlony diffusion) cannot be performed on unconcentrated supernatants, but now the new Miles monoclonal typing kits make it possible to utilize the easy-to-use Ouchterlony technique for the determination of the immunoglobulin class of mouse or rat monoclonal antibodies in tissue culture supernatants. The kits allow detection of monoclonal antibodies at down to 0.1mg per dl.

Reader Service No. 119